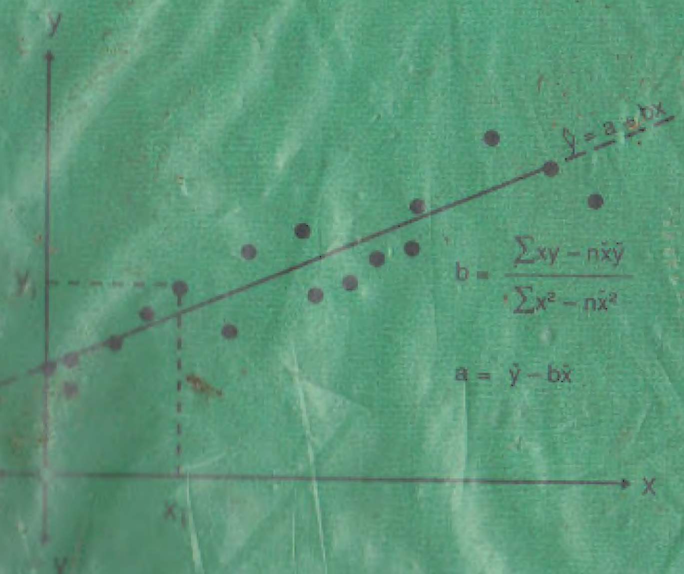
STATISTIEK VIR TECHNIKONSTUDENTE



DERDE UITGAWE 1987

P.H. KLOPPERS

1992

STATISTIEK VIR TECHNIKON STUDENTE

DERDE UITGAWE: 1987

P. H. KLOPPERS (B.Sc (HONS))

Stud no: 91015889

G. ERASMUS
Costando 205
Bourke str 17
Sunnyside
0132

Tel: (012) 3437791
~~(01536) 32786~~
(014) 763-2786

© MK UITGEWERS (BK) CK86/15751/23 (1987)
ALLE REGTE VOORBEHOU:

Geen deel van hierdie publikasie mag gereproduseer word in enige vorm hetsy elektronies, magneties, fotostaties of per bandopname sonder die skriftelike toestemming van die outeur.

INHOUD

BESKRYWENDE STATISTIEK

HOOFSTUK 1: INLEIDING

1.1 Wat is Statistiek

Enige Wetenskaplike ondersoek kom daarop neer dat gegewens (data) versamel word. Statistiek is nou die studie om sinvolle gevolgtrekkings uit hierdie data te maak.

In statistiek word metodes bestudeer om

i) data in te samel
ii) data te ontleed
iii) gevolgtrekkings uit ontledings te maak.

1.2 Verskillende tipes data

Indien waarnemings wat gemaak word slegs uit vaste geïsoleerde getalle bestaan word dit diskrete data genoem. Byvoorbeeld die ouderdom van 'n persoon in volle jare is een van (0; 1; ....; 110).

Indien waarnemings enige bepaalde waarde in 'n sekere interval kan aanneem word dit kontinue data. Byvoorbeeld die lengte van volwasse mans kan enige waarde tussen 1,0 m en 2,0 m wees. K/data is meetbaar.

Indien die waarnemings nie meetbaar is nie word daar van kwalitatiewe data gepraat. Byvoorbeeld: oogkleur kan nie gemeet word nie. Bevat 'n sekere eienskap.

In hierdie kursus bepaal ons ons slegs tot die eerste twee tipes van data en ontleed dit presies op dieselfde manier.

1.3 Sigma notasie

Gestel ons maak n metings. Dit kan voorgestel word as

$x_1;\ x_2;\ \ldots;\ x_n$.

Die som van die waarnemings word in 'n kort skryfwyse, die sigma notasie, geskryf as

$$x_1+x_2+\ldots+x_n=\sum_{i=1}^{n} x_i$$

Voorbeeld 1: Gestel ons het die ouderdomme (in jare) van 4 persone gemeet as : 8 ; 20 ; 14 en 32.

Nou is $x_1=8$ $x_2=20$ $x_3=14$ $x_4=32$ en

$$\sum_{i=1}^{4} x_i=8+20+14+32=74.$$

Eienskappe van die Sigmanotasie

1) $\sum_{i=1}^{n} x_i^2 = x_1^2 + x_2^2 + \cdots x_n^2$

2) $\left[\sum_{i=1}^{n} x_i\right]^2 = (x_1 + x_2 + \cdots + x_n)^2$

3) $\sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$

4) $\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i = (x_1 + \cdots + x_n)(y_1 + \cdots + y_n)$

5) $\sum_{i=1}^{n} k\, x_i = k \sum_{i=1}^{n} x_i$ (Let op dat k 'n konstante is.)

6) $\sum_{i=1}^{n} (x_i + y_i)$

$= \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$

7) $\sum_{i=1}^{n} k = nk$

8) $\sum_{i=1}^{n} i = 1 + 2 + \cdots + n$

$= \frac{n(n+1)}{2}$

Voorbeeld 2: Gestel $x_1=2$ $x_2=-5$ $x_3=4$ $x_4=-8$

$y_1=-3$ $y_2=-8$ $y_3=10$ $y_4=6$

dan is:

a) $\sum_{i=1}^{4} x_i = x_1+x_2+x_3+x_4 = 2-5+4-8 = -7$

b) $\sum_{i=1}^{4} y_i^2 = y_1^2+y_2^2+y_3^2+y_4^2 = (-3)^2+(-8)^2+(10)^2+(6)^2$

$= 209$

c) $\sum_{i=1}^{4} x_i y_i = (2)(-3)+(-5)(-8)+(4)(10)+(-8)(6)$

$= 26.$

OPGAWE 1

1. Gee 'n voorbeeld van 'n wetenskaplike ondersoek in u vakgebied.

2. Hoe sou u die volgende tipes data klassifiseer:

   a) 20 persone se massa word gemeet.

   b) Die leeftyd van 50 muskiete, in dae, word gemeet.

   c) Die tipe werk van 100 mense word geklassifiseer as klerklik, tegnies of bestuur.

3. Gee nog 3 voorbeelde van die verskillende tipes data.

4. Beskou die data in voorbeeld 1.3.2.

   Bereken: Antwe bl. 161

   a) $\sum_{i=1}^{4} x_i^2$ 109

   b) $\left(\sum_{i=1}^{4} x_i y_i\right)^2$ 676

   c) $\sum_{i=1}^{4} (2x_i + y_i)$ −9

   d) $\left(\sum_{i=1}^{4} x_i\right)\left(\sum_{i=1}^{4} y_i\right)$ −35

5. Beskou die data

   $x_1=4,2$ $x_2=6,4$ $x_3=3,6$ $x_4=3,1$ $x_5=5,3$

   $y_1=10,7$ $y_2=11,3$ $y_3=9,1$ $y_4=13,0$ $y_5=8,4$ $y_6=12,2$

   Bereken:

   a) $\sum_{i=1}^{4} x_i y_i$ 190,32

   b) $\sum_{i=1}^{3} x_i^3$ 382.88

   c) $\left(\sum_{i=1}^{5} y_i\right)^2$ 2756,25

   d) $\sum_{i=1}^{5} x_i \cdot \sum_{i=1}^{5} y_i$ 1086.5

   e) toon aan dat $\sum_{i=1}^{3} x_i y_i \neq \sum_{i=1}^{3} x_i \sum_{i=1}^{3} y_i$

   en dat $\sum_{i=1}^{3} x_i^2 \neq \left[\sum_{i=1}^{3} x_i\right]^2$.

HOOFSTUK 2: STEEKPROEFNEMING

2.1 Inleiding

As 'n studie op 'n groep items gedoen moet word dan noem ons die versameling wat al hierdie items insluit die populasie. Byvoorbeeld die populasie van motorkarre in S.A., die populasie van blanke mans wat in die Transvaal woon, die populasie van al die beeste in S.A. wat aan hartwater ly.

'n Steekproef word gedefinieer as enige deelversameling van 'n populasie.

Daar word baie selde 'n ondersoek op die hele populasie (sensus) gedoen want:

a) dit is duur
b) neem te lank
c) populasie is te groot
d) eenhede word vernietig by ondersoek.
e) sensus kan baie onakkuraat raak.

Dus word 'n studie normaalweg beperk tot 'n steekproef uit die populasie.

2.2 Metodes van Steekproefneming

Dit spreek vanself dat daar baie versigtig te werk gegaan moet word wanneer 'n steekproef geneem word, aangesien dit nou slegs die steekproefelemente is wat aan die ondersoek onderwerp gaan word terwyl inligting aangaande die populasie as geheel verlang word. Dit is dus uiters belangrik dat die steekproef verteenwoordigend van die populasie moet wees.

2.2.1 Eenvoudige ewekansige steekproefneming Simple random sampling

Hier word die steekproef so geneem dat elke element in die populasie 'n ewe groot kans het om in die steekproef opgeneem te word.

Gestel daar is N elemente in die populasie en ons wil 'n steekproef van n elemente trek. Nommer die populasie van 1 tot N. Met behulp van die kansgetalletabel, Tabel XX van Stoker, word n getalle tussen 1 en N getrek en die ooreenstemmende items word in die steekproef opgeneem.

Voorbeeld 1

Gestel ons het 'n populasie van 50 mense en wil 'n steekproef van 10 mense trek en hulle lengtes meet. Nommer die populasie. Uit Tabel XX vanaf kol 11 ry 11 lees ons nou die tien kansgetalle af:

33 03 46 48 11 28 39 25 09 41

Meet nou die lengtes van hierdie tien persone.

(Begin by 'n willekeurige plek in die tabel. Neem genoeg syfers in die kansgetal op. Beweeg af of dwars in die tabel. Laat kansgetalle wat te groot is net weg.)

2.2.2 Gestratifiseerde ewekansige steekproefneming

Hierdie metode word gebruik wanneer die populasie se elemente baie verskillend (heterogeen) is met betrekking tot die eienskap wat ons wil meet.

Eers word die populasie in homogene groepe (strata) verdeel. Nou word uit elke stratum 'n eenvoudige ewekansige steekproef geneem.

Gestel $N=\sum_{i=1}^{k} N_i$ en $n=\sum_{i=1}^{k} n_i$, met N die totale aantal in die populasie, N_i die aantal in die ide sub populasie, en n_i die steekproef grootte uit die ide sub populasies, dan is $n_j=\frac{N_j}{N}n$.

Voorbeeld 2:

Uit 100 persone wil ons 'n steekproef van 20 persone neem en die massas bestudeer.

Om weet dat die massas van mans en dames is nie dieselfde nie.

Ons moet dus die 100 persone in 2 strata, mans en dames, verdeel.

Gestel daar is 40 dames en 60 mans.

Nou is $N=100$ $n=20$ $k=2$ $N_1=40$ $N_2=60$

en $n_1=\frac{40}{100}\cdot 20=8$

$n_2=\frac{60}{100}\cdot 20=12$

2.2.3 Sistematiese steekproefneming

Hier word die elemente van die populasie genommer.

Een kansgetal tussen 1 en $\frac{N}{n}$, sê r., word getrek.

Die steekproef is nou die elemente met nommers

r ; $r+(\frac{N}{n})$; $r+(2\frac{N}{n})$; $r+(3\frac{N}{n})$; .... .

N = populasie
n = aantal

n = die element/syfer in s.proef.

Voorbeeld 3

Gestel N=500 en ons wil 'n sistematiese steekproef van grootte n=50 trek.

$\frac{N}{n}=\frac{500}{50}=10$

Trek een kansgetal tussen 1 en 10 sê 6. Nou word die populasie elemente met nommers.

6 ; 16 ; 26 ; 36 ; .... ; 496 in die steekproef opgeneem.

2.3 Waar 'n steekproef op een van bogenoemde metodes getrek word kan aangeneem word dat dit dieselfde eienskappe as die populasie sal besit. (Die gemiddelde waarde uit die steekproef sal byvoorbeeld baie naby wees aan die gemiddelde waarde uit die populasie.)

OPGAWE 2

1. Gestel N=1000 en n=20. (Sien datastel bylae 2.)

   a) beplan 'n eenvoudige ewekansige steekproef

   b) beplan 'n sistematiese steekproef

   c) gestel ons kan die elemente van die populasie in 4 groepe verdeel waar $N_1=200$; $N_2=150$; $N_3=50$; $N_4=100$. Beplan 'n gestratifiseerde ewekansige steekproef.

2. Definieer drie situasies uit u vakgebied waar die drie steekproef metodes goeie toepassings sal hê.

3. 'n Steekproef van 1000 motoriste in Pretoria moet geneem word. Ontwerp 'n geskikte steekproefplan.



HOOFSTUK 3

DIE ONTLEDING VAN ONGEGROEPEERDE DATA

3.1 Algemeen

Gestel ons het 'n steekproef van grootte n uit 'n populasie getrek en 'n sekere eienskap gemeet. Die data kan ons voorstel as

$x_1 ; x_2 ; \ldots ; x_n$

Sulke data noem ons ongegroepeerd.

Voorbeeld 1

By die Buitepasiënte afdeling van 'n groot hospitaal is 'n steekproef van tien pasiënte geneem en die wagtyd (in minute) totdat hulle behandeling ontvang was asvolg:

35 22 16 31 23 47 108 35 39 10

(hier is $x_1=35$; $x_2=22$; ... ; $x_{10}=10$)

3.2 Maatstawwe van posisie: Ongegroepeerde Data

Die doel van hierdie maatstawwe is om 'n aanduiding te kry van die sentraliteitspunt van die data. (sien ook § 5)

3.2.1.1 Die Rekenkundige gemiddeld ($\bar{x}$)

Die rekenkundige gemiddeld van 'n stel ongegroepeerde data is die som van die waarnemings gedeel deur die aantal waarnemings:

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i$$

Vir die data in voorbeeld 3.1.1 is

$$\bar{x}=\frac{1}{10}(366)=36,6$$

3.2.1.2 Die Meetkundige gemiddeld (Mg)

$$Mg = \sqrt[n]{x_1 \cdot x_2 \cdot \cdots x_n}$$

Die meetkundige gemiddeld word gewoonlik gebruik as die waarnemings verhoudings van getalle is.

Voorbeeld 2

Die groei van die bates van 'n sekere firma vir die afgelope 3 jaar, uitgedruk as 'n verhouding tot die vorige jaar se bates was $\frac{6}{4}$ $\frac{8}{6}$ en $\frac{9}{8}$. Bereken die gemiddelde groei vir die afgelope 3 jaar.

Oplossing •= maal met mekaar.

$$Mg = \sqrt[3]{\frac{6}{4} \cdot \frac{8}{6} \cdot \frac{9}{8}} = 1,31 = \frac{131}{100}$$ (Let op $\bar{x}$ = 1,319)

3.2.1.3 Die Harmoniese gemiddeld (Hg)

$$Hg = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

Die harmoniese gemiddeld word gewoonlik gebruik as die waarnemings uitgedruk word per meetingseenheid.

Voorbeeld 3

Drie persone neem onderskeidelik 3, 5 en 2 uur om 'n sekere stuk werk te doen. Hoe lank sal dit hulle gemiddeld neem om die taak saam te doen?

Oplossing

$$Hg = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{3}{\frac{1}{3}+\frac{1}{5}+\frac{1}{2}} = 2,9 \text{ uur}$$

(Vergelyk dit met $\bar{x}$ = 3,33 uur.)

3.2.2 Die Mediaan (Me)

Die mediaan van 'n stel ongegroepeerde data word gedefinieer as die $\frac{n+1}{2}$ de waarde as die data van klein na groot geskik is.

Vir die data in voorbeeld 3.1.1 is

$\frac{n+1}{2} = \frac{10+1}{2} = 5,5$. Die gerangskikte data is:

10 16 22 23 31 35 35 39 47 108. En

$Me=\frac{31+35}{2} = \frac{66}{2} = 33.$

(Let op dat as n onewe is kan ons presies die middelste waarde neerskryf.)

3.2.3 Die Modus (Mo)

Die modus van 'n stel ongegroepeerde data word gedefinieer as die waarde wat die meeste kere voorkom. In sommige gevalle mag daar meer as een modus wees en in ander gevalle mag daar glad nie 'n modus wees nie.

Vir die data in voorbeeld 3.1.1 is

Mo=35.

3.2.4 Vergelyking tussen $\bar{x}$, Me en Mo

$\bar{x}$ is maklik om te bereken en eenvoudig om te begryp. Dit is op elke getal in die data gebaseer en is dus verteenwoordigend. $\bar{x}$ is egter baie afhanklik van foutiewe waardes wat baie groot of baie klein is. (uitskieters)

Me is ook maklik berekenbaar en die betekenis daarvan kan maklik verstaan word. Dit is minder afhanklik van uitskieter waardes. 'n Nadeel van Me is dat dit slegs deur 1 of 2 waardes in die data bepaal word.

Mo is ook maklik berekenbaar en verstaanbaar. In sommige gevalle lewer $\bar{x}$ 'n onsinnige grootheid terwyl Mo sinvol is (dink aan skoengroottes).

3.2.5 Kwantiele

Die drie kwartiele, q_1, q_2 en q_3 verdeel die geskikte data in vier gelyke dele.

Die nege desiele, d_1, d_2, ..., d_9 verdeel die geskikte data in tien gelyke dele.

Die nege-en-negentig persentiele verdeel die geskikte data in honderd gelyke dele.

Let op dat:

$Me = q_2 = d_5 = P_{50}$

$q_1 = P_{25} \quad q_3 = P_{75}$

$d_1 = P_{10} \quad d_2 = P_{20}$ ens.

Dit is dus slegs nodig om die persentiele te kan bereken.

Die i-de persentiel word gedefinieer as die $(i(n+1))/100$ste waarneming in die skikking.

Voorbeeld 4

Vir die data in voorbeeld 3.1.1 is die skikking:

10 16 22 23 31 35 35 39 47 108

i. Vir $P_{40} = d_4$: $\frac{i(n+1)}{100} = \frac{40(11)}{100} = 4,4$

Dus is $P_{40} = d_4 = 23 + 0,4(31-23) = 26,2.$

ii. Vir $q_3 = P_{75}$: $\frac{i(n+1)}{100} = \frac{75(11)}{100} = 8,25$

Nou is $q_3 = 39 + 0,25(47-39) = 41,0.$

3.3 Maatstawwe van Verspreiding: Ongegroepeerde Data

Beskou die twee stelle data:

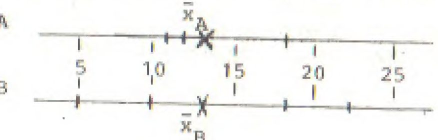
A: 13 11 18 14

B: 18 5 23 10

$\bar{x}_A = \frac{1}{4}(56) = 14 \qquad \bar{x}_B = \frac{1}{4}(56) = 14$

Dit lyk asof die twee stelle data dieselfde is.

Grafies kan ons nou die data asvolg voorstel:



Hieruit sien ons dat alhoewel $\bar{x}_A$ en $\bar{x}_B$ dieselfde is lê die data van B baie meer verspreid as die data van A.

Dus is lokaliteitsmaatstawwe nie voldoende om 'n stel data te beskryf nie. Ons moet ook maatstawwe van verspreiding definieer.

3.3.1 Die Variasiebreedte (R)

R = grootste waarde - kleinste waarde.

Vir die data in voorbeeld 3.1.1 is

R = 108-10 = 98.

R is nie 'n baie geskikte maat van die verspreiding nie aangesien dit slegs deur twee van die waardes bepaal word.

3.3.2 Die gemiddelde afwyking (G)

Indien ons 'n maatstaf van spreiding wil definieer wat op al die waardes in die datastel gebaseer is, moet 'n verwysingspunt in die datastel gekies word en die afwykings van al die waardes van dié verwysingspunt moet dan bereken word. Die mees geskikte verwysingspunt in 'n datastel is die rekenkundige gemiddeld $\bar{x}$.

$\bar{x} \leftarrow x_i - \bar{x} \rightarrow x_i$

ŉ Logiese definisie vir ŉ spreidingsmaatstaf is dus:

$$\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x}).$$

Hierdie maatstaf is egter nie bruikbaar nie aangesien positiewe afwykings en negatiewe afwykings mekaar uitkanselleer. Trouens, dit kan bewys word dat

$\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})=0$ vir enige datastel.

Die oplossing van die probleem lê daarin dat die tekens van die verskille nie in ag geneem moet word nie. Dit lei dan tot die onderstaande definisie van die gemiddelde afwyking:

$$G=\frac{1}{n}\sum_{i=1}^{n}|x_i-\bar{x}|$$

Beskou weer die data in voorbeeld 3.1.1.

$\bar{x}=36,6$

Ons stel die volgende berekeningstabel op:

TABEL 1

x	$(x-\bar{x})$	$\lvert x-\bar{x}\rvert$
35	- 1,6	1,6
22	-14,6	14,6
16	-20,6	20,6
31	- 5,6	5,6
23	-13,6	13,6
47	10,4	10,4
108	71,4	71,4
35	- 1,6	1,6
39	2,4	2,4
10	-26,6	26,6
366		168,4

Nou is $G=\frac{1}{n}\sum_{i=1}^{n}|x_i-\bar{x}|$

$=\frac{1}{10}(168,4)=16,84.$

ŉ Probleem met die gemiddelde afwyking is die neem van die absolute waarde. Dit is wiskundig nie heeltemal korrek nie en alhoewel G ŉ baie goeie maat van verspreiding is kan ons dit nie vir verdere analise gebruik nie.

3.3.3 Die Variansie en Standaardafwyking (S^2 en S)

ŉ Ander manier om van die negatiewe waardes by $(x_i-\bar{x})$ ontslae te raak is om die waardes te kwadreer, $(x_i-\bar{x})^2$.

Nou word die variansie gedefinieer as

$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$ (definisie formule) moet nie gebruik nie.

$=\frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2-n\bar{x}^2\right)$ (berekenings formule)

$S=\sqrt{S^2}=\sqrt{\frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2-n\bar{x}^2\right)}$ noem ons die standaardafwyking.

Beskou weer die data in voorbeeld 3.1.1 en Tabel 1.

Om S^2 en S te bereken voeg ons nog 1 kolom by Tabel 1 nl. x^2.

Tabel 2

x	$(x-\bar{x})$	$\lvert x-\bar{x}\rvert$	x^2
35	- 1,6	1,6	1225
22	-14,6	14,6	484
16	-20,6	20,6	256
31	- 5,6	5,6	961
23	-13,6	13,6	529
47	10,4	10,4	2209
108	71,4	71,4	11664
35	- 1,6	1,6	1225
39	2,4	2,4	1521
10	-26,6	26,6	100
			20174

$S^2=\frac{1}{n-1}(\Sigma x_i^2-n\bar{x}^2)$

$=\frac{1}{9}(20174-10(36,6)^2)$

$=753,15$

$S=\sqrt{S^2}$

$=27,44$

3.3.4 Die Koëffisiënt van Variasie V

Wanneer verspreidingsmaatstawwe van twee of meer datastelle met mekaar vergelyk wil word, en die rekenkundige gemiddeldes van die datastelle verskil baie van mekaar, kan so 'n vergelyking slegs getref word indien die koëffisiënt van variasie vir elke datastel bereken word.

Definisie:

$V=\frac{S}{\bar{x}}x100\%$

Vir die data van voorbeeld 3.1.1 is

$V=\frac{27,44}{36,6}$ x 100 = 74,97%

Hoe kleiner V hoe meer betroubaar is die steekproef data. 'n Baie growwe reël is dat V kleiner moet wees as 20% anders is die steekproef nie baie goed nie.

OPGAWE 3

1. Die volgende data is die bakterie-tellings in 12 biologiese eksperimente:

138 146 168 146 161 164 158 126 173 145 150 140

Bereken

i) $\bar{x}$, Mg, Hg
ii) Mo
iii) Me
iv) R
v) G
vi) S² en S
vii) V

Interpreteer ook elke grootheid.

2. Beskou die volgende twee stelle data:

A : 19,0 10,0 12,8 25,8 19,9 12,0 22,48
B : 50,0 33,0 27,1 37,5 29,0 40,4

Bereken V_A en V_B en vergelyk die groothede.

3. Die massas van 10 skaaplammers, op 6 maande ouderdom was as volg:

11,9 10,3 13,5 12,4 11,9 9,8 15,0 13,2 8,9 10,4

Bereken i) $\bar{x}$, Mo, Me, Mg, Hg

ii) G, S², V

4. 'n Gestratifiseerde ewekansige steekproef het die volgende opbrengste gehad:

Stratum	data								
1	31,1	34,3	39,4	42,8	47,2	56,3	62,4	66,0	
2	32,2	36,5	38,7	41,2	48,9	59,9	73,5	84,8	91,1
3	28,3	31,6	35,0	37,1	38,6	41,1	46,7		

i) Vir elke stratum bereken $\bar{x}$ en S².

ii) Bereken die algehele gemiddelde en variansie.

5. Doen die berekenings van opgawes 3.3 en 3.4 met behulp van 'n sakrekenaar.

6. Bereken Mg en Hg vir die data ...

HOOFSTUK 4

TABELLERING EN GRAFIESE VOORSTELLING VAN GEGEWENS

4.1 Frekwensieverdelings

Die ontledingsmetodes wat in hoofstuk 3 bespreek is werk baie moeilik as ŉ baie groot raak ($n>50$). In hierdie geval is dit makliker om die data te tabuleer en dan die berekeninge te maak. Dit word gedoen deur die sogenaamde frekwensietabel op te stel.

Uit die tabelle is dit ook moontlik om die data grafies voor te stel ten einde ŉ duideliker beeld van die data te kry.

ŉ Frekwensietabel bestaan uit kategorieë of klasse wat gedefinieer is en waarin elke waarneming geklassifiseer kan word.

Voorbeeld 1

Die wagtyd (in minute) van 50 buitepasiënte by ŉ sekere hospitaal voordat hulle behandeling ontvang het, is soos volg:

35	22	63	6	49	19	15	83	46	19
16	31	24	29	36	68	42	57	64	8
23	47	21	51	7	40	19	46	16	32
108	33	55	32	22	36	25	27	37	58
39	10	42	28	72	13	51	45	77	16

Hierdie data kan as volg geklassifiseer word:

TABEL 3

Wagtyd in minute (klasse)	Aantal pasiënte (Frekwensie)
6 en minder as 21	12
21 en minder as 36	14
36 en minder as 51	12
51 en minder as 66	7
66 en minder as 81	3
81 en minder as 96	1
96 en minder as 111	1
TOTAAL	50

Om skryfwerk te vergemaklik, word die klasse en dus die frekwensieverdeling soos volg weergegee:

TABEL 4

Frekwensieverdeling van die wagtyd (in minute) van 50 buitepasiënte by ŉ bepaalde hospitaal

Wagtyd in minute	Frekwensie (f)
[6 ; 21)	12
[21 ; 36)	14
[36 ; 51)	12
[51 ; 66)	7
[66 ; 81)	3
[81 ; 96)	1
[96 ; 111)	1
TOTAAL	50

Let op dat die klasse só gekies word dat elke waarneming in een en slegs een klas kan val.

4.2 Die konstruksie van frekwensieverdelings

Die volgende stappe dien as riglyne vir die opstel van ŉ frekwensieverdeling.

i) Bereken die variasiewydte

R = grootste waarde − kleinste waarde

ii) Verkry ŉ aanduiding van die aantal klasse wat opgestel moet word, deur gebruik te maak van die reël van Sturges.

Aantal klasse $k = 1+3{,}3\log_{10} n$, afgerond tot ŉ gerieflike getal.

iii) Verkry ŉ aanduiding van die klaswydte:

klaswydte $c=\frac{R}{k}$, afgerond tot ŉ gerieflike getal.

iv) Stel die klasse op, sodanig dat geen twee klasse oorvleuel nie en sodat die kleinste waarde in die eerste klas lê en die grootste in die laaste klas.

v) Tel die aantal waardes wat in elke klas val di. bereken die frekwensies van die klasse.

Voorbeeld 2

Beskou weer voorbeeld 4.1.1.

i) Variasiebreedte R=108-6=102

ii) k=1+3,3log50=6,6 ≃7

iii) $c=\frac{R}{k}=\frac{102}{7}=14,5 \quad \simeq 15$

iv) Die klasse is dus:

[6 ; 21) [21 ; 36) [36 ; 51) [51 ; 66)
[66 ; 81) [81 ; 96) [96 ; 111)

v) Om die berekening van die frekwensies te vergemaklik gebruik ons 'n telkolom:

TABEL 5

Frekwensieverdeling

v)

Klasse	Telling	Frekwensie
[6 ; 21)	𝍸 𝍸 ǁ	12
[21 ; 36)	𝍸 𝍸 ǁǁ	14
[36 ; 51)	𝍸 𝍸 ǁ	12
[51 ; 66)	𝍸 ǁ	7
[66 ; 81)	ǀǀǀ	3
[81 ; 96)	ǀ	1
[96 ; 111)	ǀ	1
TOTAAL		50 Σ

NB Let op dat $\sum_{i=1}^{k} f_i = n$ — $F_i = n$

4.3 Kumulatiewe frekwensieverdelings

Kumulatiewe frekwensieverdeling word uit die frekwensieverdeling verkry, deur die frekwensies agtereenvolgens bymekaar te tel of te akkumuleer.

Vir die data in voorbeeld 4.1.1 kan ons die frekwensietabel as volg uitbrei:

TABEL 5

Klas	f	Kumulatiewe frekwensie F
[6 ; 21)	12	12
[21 ; 36)	14	26
[36 ; 51)	12	38
[51 ; 66)	7	45
[66 ; 81)	3	48
[81 ; 96)	1	49
[96 ; 111)	1	50

n=50

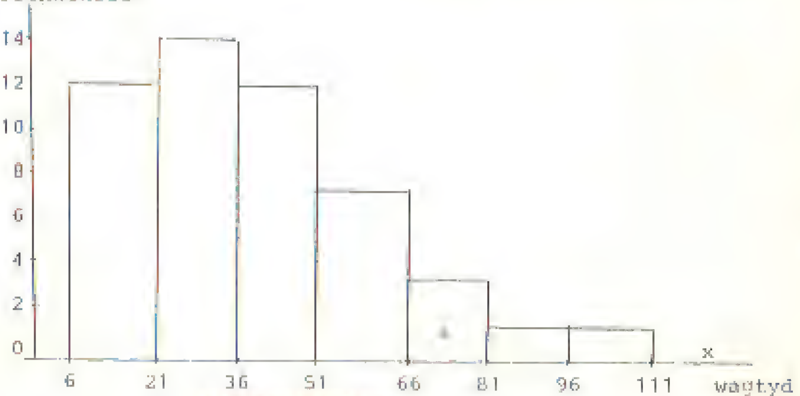
4.4 Grafiese voorstellings

Frekwensieverdelings kan op twee maniere grafies voorgestel word, naamlik deur middel van 'n histogram en 'n frekwensieveelhoek.

4.4.1 Die histogram

'n Histogram word gekonstrueer deur die klasse (intervalle) op die x-as uit te stip en die frekwensies op die y-as uit te stip. Reghoekies word dan geteken op die assestelsel, waarby die lengtes van die reghoekies gelyk is aan die klasfrekwensies en die breedtes gelyk is aan die klaswydte. Beskou weer voorbeeld 4.1.1.

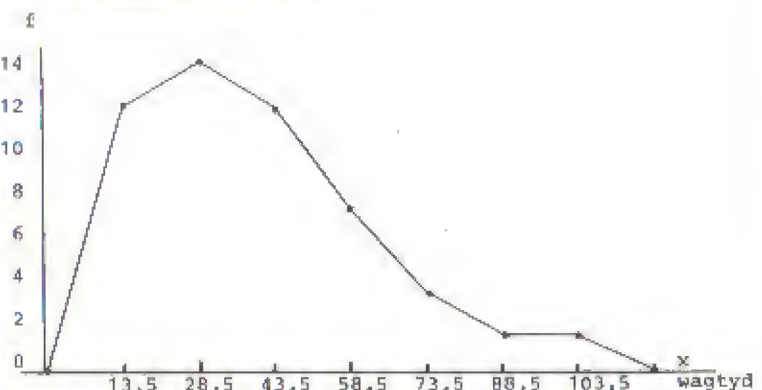
Histogram van die wagtyd (in minute) van 50 buitepasiënte by 'n bepaalde hospitaal.



4.4.2 Die Frekwensieveelhoek

In hierdie geval word die middelwaardes van die klasse op die x-as uitgestip en die frekwensies op die y-as. Die punte op die grafiek word dan met reguitlyne verbind. Beskou weer die voorbeeld 4.4.1.

Frekwensieveelhoek van die wagtyd van 50 buitepasiënte by 'n bepaalde hospitaal.



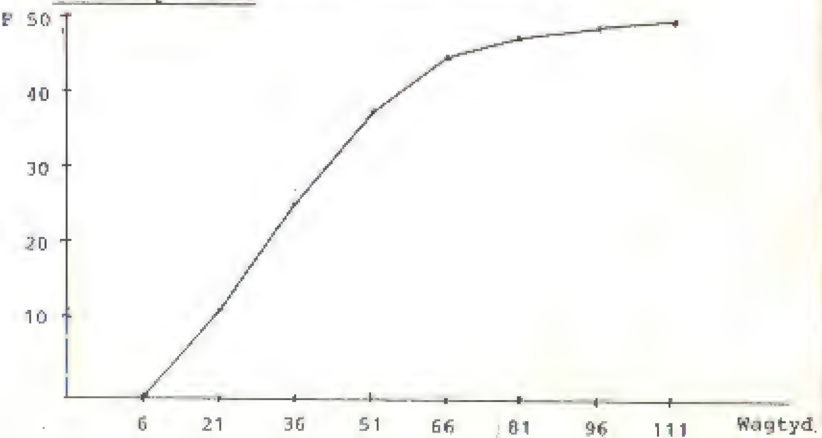
Twee denkbeeldige middelwaardes, een aan die linkerkant en een aan die regterkant, word bygevoeg sodat die figuur op die x-as kan sluit.

4.4.3 Die Kumulatiewe frekwensieveelhoek

Hier word die kumulatiewe frekwensie uitgestip teenoor die bo klasgrens, en die veelhoek word aan die linkerkant met die x-as verbind.

Vir die data van voorbeeld 1:

Kumulatiewe frekwensieveelhoek van die wagtye van 50 buitepasiënte



4.5 Sektorkaarte

'n Sektorkaart bestaan uit 'n sirkel wat in verskillende sektore verdeel is sodat die oppervlaktes van die sektore eweredig is aan die verhouding data items wat tot 'n sekere groep behoort.

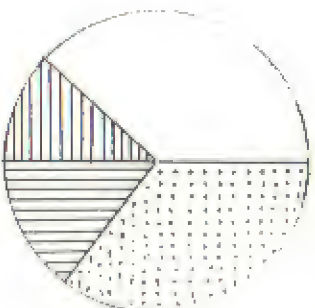
Die grootte van die hoek van 'n sektor = (Verhouding waarnemings) x 360°.

Voorbeeld 4

Die vermoëns van 180 persone om 'n sekere taak te verrig is gemeet en as volg geklassifiseer:

klas	aantal persone	verhouding	hoek (°)
Baie goed	21	0,117	42,12
goed	70	0,389	140,04
swak	63	0,35	126,0
baie swak	26	0,144	51,84
	180	1	360

Hierdie data kan as volg in 'n sektorkaart voorgestel word:



4.6 Staafkaarte

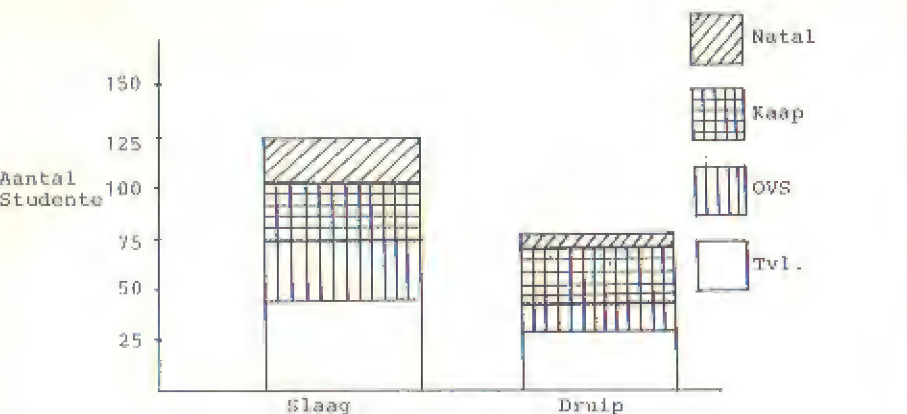
Die verhoudings wat vir sektorkaarte bereken is kan ook gebruik word om 'n staaf in dele te verdeel.

Voorbeeld 5

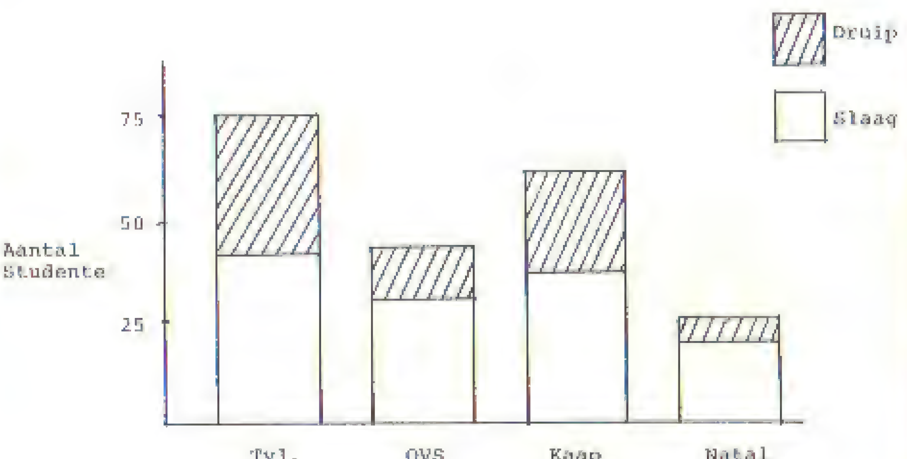
Die volgende tabel gee die persentasies van 200 studente in 'n sekere statistiek toets:

	Tvl	OVS	Kaap	Natal	
Slaag	45	29	33	18	125
Druip	30	11	27	7	75
	75	40	60	25	200

Hierdie data kan as volg in 'n staafkaart voorgestel word.



OF



OPGAWE 4

1. Stel 'n frekwensieverdeling op vir die volgende datastel.

138	150	144	149
146	140	136	152
168	138	163	154
146	142	135	140
161	135	150	145
164	132	125	157
150	147	148	144
126	176	110	165
173	147	153	135
145	142	156	128

2. Onderstaande tabel verteenwoordig die weeklikse oortydbetaling (in Rand) van 40 werknemers by die XYZ-fabriek. Stel 'n frekwensieverdeling op.

42,15	62,30	68,80	55,60
67.25	43,90	45,70	58,40
53,70	57,00	43,20	63,40
59,80	48,80	37,20	80,50
40,10	69,30	65,10	56,40
47,80	30,10	62,40	79,10
71,50	58,60	42,60	43,35
57,90	49,30	59,30	56,65
41,20	55,50	42,40	67,60
63,50	78,40	48,20	63,35

3. Onderstaande tabel verteenwoordig die belasting betaal (in R1000) deur 'n steekproef van 50 maatskappye in 'n bepaalde ekonomiese sektor gedurende die 1980/81 belastingjaar. Stel 'n frekwensieverdeling op.

27,05	21,05	20,01	15,51	15,51
24,04	19,25	13,55	40,00	19,01
30,00	6,50	21,05	42,25	32,04
57,00	60,00	28,00	49,61	18,31
30,00	15,75	17,65	25,00	22,40
16,05	4,50	43,00	33,55	18,35
30,00	13,85	40,00	28,00	20,00
19,55	37,35	20,81	22,58	30,01
36,74	26,55	28,10	38,00	15,31
28,45	35,50	27,10	25,00	36,60

4. Skets die histogramme en frekwensieveelhoeke van opgawes 4.1, 4.2 en 4.3.

5. Die resultate wat 100 studente in 'n Wiskundetoets behaal het was as volg:

%	0-20	21-40	41-50	51-70	70+
Aantal Studente	11	17	42	21	9

Stel die data op 'n sektorkaart voor.

6. Die eise wat 'n sekere versekeringsmaatskappy vir 'n maand gehad het kan as volg opgesom word:

		Waarde van eis	
		tot R1000	bo R1000
Ouderdom van eiser	18-25	43	16
	26-35	10	36
	36+	4	9

Stel die data in 'n staafkaart voor.

7. Stel die kumulatiewe frekwensieverdeling op vir die oefeninge in opgawes 4.1 tot 4.3.

(8.) Stel die kumulatiewe frekwensie op vir die volgende frekwensieverdeling.

(a)

Klasse	Frekwensie (f_i)	F_i
[0 ; 2)	94	94
[2 ; 4)	107	201
[4 ; 6)	70	271
[6 ; 8)	62	333
[8 ; 10)	31	364
[10 ; 12)	28	392

(b)

Klasse	Frekwensie (f_i)	F_i
[30,00 ; 40,00)	4 (is kont data)	4
[40,00 ; 50,00)	16	20
[50,00 ; 60,00)	13	33
[60,00 ; 70,00)	12	45
[70,00 ; 80,00)	6	51
[80,00 ; 90,00)	2	53

(9.) Onderstaande frekwensietabel verteenwoordig 'n steekproef van 1827 rokers, geklassifiseer volgens ouderdom, wat simptome van 'n sekere longdefek toon. Stel die kumulatiewe frekwensieverdeling op.

Ouderdom	Frekwensie (f_i)	F_i
[20 ; 25)	9	9
[25 ; 30)	23	32
[30 ; 35)	54	86
[35 ; 40)	121	207
[40 ; 45)	169	376
[45 ; 50)	269	645
[50 ; 55)	404	1041
[55 ; 60)	406	1455
[60 ; 65)	372	1827

10. Skets die frekwensiehistogramme en die frekwensieveelhoeke vir die opgawes 4.8 en 4.9.

11. Skets die kumulatiewe frekwensieveelhoeke van opgawes 4.8 en 4.9.

HOOFSTUK 5

DIE ONTLEDING VAN GEGROEPEERDE DATA

5.1 Algemeen

By 'n stel gegroepeerde data (sien tabel 5 bls. 17) bereken ons presies dieselfde groothede as by ongegroepeerde data, ons maak net van ander formules gebruik. Die interpretasie van die groothede is ook presies dieselfde as by ongegroepeerde data.

Ons brei die frekwensie tabel asvolg uit: (sien tabel 5 bls. 17)

TABEL 6

Klas	f	F	x	fx	fx²
[6 ; 21)	12	12	13,5	162,0	2187,00
[21 ; 36)	14	26	28,5	399,0	11371,50
[36 ; 51)	12	38	43,5	522,0	22707,00
[51 ; 66)	7	45	58,5	409,5	23955,75
[66 ; 81)	3	48	73,5	220,5	16206,75
[81 ; 96)	1	49	88,5	88,5	7832,25
[96 ; 111)	1	50	103,5	103,5	10712,25
	n=50			$\sum_{i=1}^{k} f_i x_i = 1905,0$	$\sum_{i=1}^{k} f_i x_i^2 = 94972,5$

x is die waarde in die middel van die klas (klasmiddelwaarde).

k is die aantal klasse.

5.2 Maatstawwe van posisie: Gegroepeerde Data

5.2.1 Die Rekenkundige gemiddeld ($\bar{x}$)

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{k} f_i x_i$$

Vir tabel 6:

$$\bar{x} = \frac{1}{50} (1905,0) = 38,1$$

5.2.2 Die Modus (Mo)

$$Mo = L + C \left(\frac{f_m - f_{m-1}}{2f_m - f_{m-1} - f_{m+1}}\right)$$

Die modaleklas is die klas met die grootste frekwensie
L is die onderste klasgrens van die modaleklas.
C is die klaslengte.
f_m is die frekwensie van die modaleklas.

f_{m-1} is die frekwensie van die klas voor die modaleklas.
f_{m+1} is die frekwensie van die klas na die modaleklas.

Vir tabel 6:

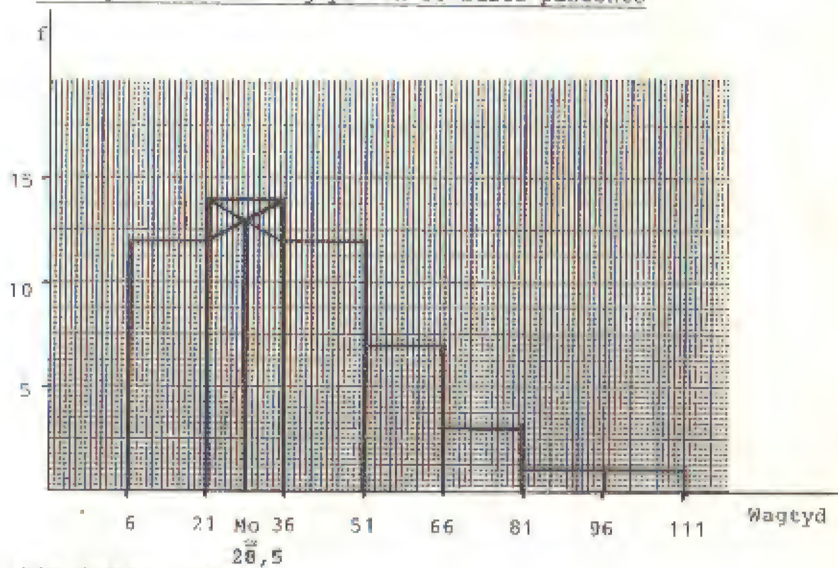
Die Modaleklas is die klas [21 ; 36).

$$Mo = 21 + 15 \left(\frac{14-12}{(2)(14)-12-12}\right)$$
$$= 21 + 15 (0,5)$$
$$= 21 + 7,5$$
$$= 28,5$$

5.2.2.1 Grafiese bepaling van die Modus

Die modus kan as volg uit die histogram bepaal word:

Histogram van die wagtye van 50 buite pasiënte



5.2.3 Die Mediaan (Me)

Om die mediaanklas te bepaal bereken ons $\frac{n}{2}$ en gaan af in die F kolom tot by die eerste klas waar $F \geq \frac{n}{2}$.

$$Me = L + C \left(\frac{\frac{n}{2} - F_{-1}}{f_{me}}\right)$$

L is die onderste klasgrens van die mediaanklas.
C is die klas lengte.
n is die aantal waarnemings.
f_{me} is die frekwensie van die mediaanklas.
F_{-1} is die kumulatiewefrekwensie van die klas voor die mediaanklas.

Vir tabel 6:

$\frac{n}{2} = \frac{50}{2} = 25$

Die Me-klas is die klas [21 ; 36).

$$Me = 21 + 15 \left[\frac{\frac{50}{2} - 12}{14}\right]$$

$= 21 + 13,93$

$= 34,93$

5.2.4 Kwantiele

Net soos by ongegroepeerde data kan ons ook kwartiele, desiele en persentiele bereken.

Die i-de persentiel word bereken as:

$$P_i = L + C\left(\frac{\frac{in}{100} - F_{-1}}{f_p}\right)$$

waar

L die onderste klasgrens van die i-de persentiel klas is. (Bereken $\frac{in}{100}$ gaan in F kolom af tot by eerste klas waar $\frac{in}{100} \leqslant F$).

C die klaslengte is.

F_{-1} die kumulatiewe frekwensie van die klas voor dié persentiel klas is.

f_p die frekwensie van dié persentiel klas is.

Gestel ons wil $d_7 = p_{70}$ bereken in die voorbeeld:

$$\frac{in}{100} = \frac{70(50)}{100} = 35$$

Die d_7-de klas is die klas [36;51)

$D_7 = 36+15\left(\frac{35-26}{12}\right)$

$= 36 + 11,25$

$= 47,25$

5.2.4.1 Grafiese bepaling van kwantiele

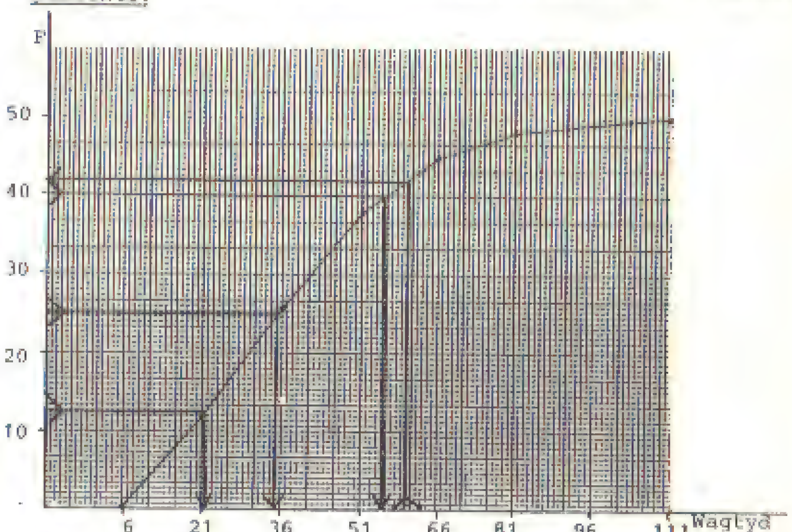
Kwantiele kan as volg uit die kumulatiewe frekwensieveelhoek bepaal word:

Vir q_1: $25n/100 = 25(50)/100 = 12,5$

Me: $50n/100 = 50(50)/100 = 25$

d_8: $80n/100 = 80(50)/100 = 40$

Kumulatiewe frekwensieveelhoek van die wagtye van 50 buite pasiënte.



$q_1 \simeq 21,5$ Me $\simeq 35$ $d_8 \simeq 54,5$

Ongeveer 42 of $\frac{42}{50} \times 100 = 84\%$ persone het minder as 60 minute gewag.

5.3 Maatstawwe van verspreiding: Gegroepeerde Data

5.3.1 Die Variansie en Standaardafwyking (S^2 en S)

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{k} f_i x_i^2 - n\bar{x}^2\right)$$

Vir tabel 6:

$S^2 = \frac{1}{49}(94972,5 - 50(38,1)^2)$

$= \frac{1}{49}(22392,2)$

$= 456,98$

$S = \sqrt{S^2}$

$= 21,38$

5.3.2 Die Koëffisiënt van Variasie (V)

$V = \frac{S}{\bar{x}} \times 100\ \%$

Vir tabel 6:

$V = \frac{21,38}{38,1} \times 100$

$= 56,12\ \%$

5.4 Mate van Simmetrie en Kurtose

5.4.1 Simmetrie (Skeefheid)

Beskou ons weer die frekwensieveelhoek van 'n verdeling sien ons dat die vorm van die kromme een van die volgende kan wees:



Ons sê die eerste verdeling is simmetries, die tweede negatief skeef en die derde positief skeef.

Vir 'n negatiefskewe verdeling geld data

$\bar{x} < Me < Mo$

en vir 'n positiefskewe verdeling geld data

$\bar{x} > Me > Mo$

'n Maat wat die skeefheid van 'n verdeling meet is Pearson se skeefheidkoëffisiënt:

$sk(P) = \frac{\bar{x} - Mo}{S}$

Indien sk(P) > 0 is die verdeling positief skeef en as sk(P) < 0 is die verdeling negatief skeef.

Vir die voorbeeld is

$sk(P) = \frac{38,1 - 28,5}{21,38} = 0,449$

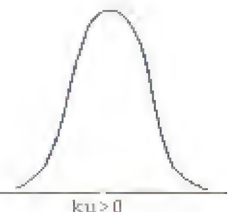
Dan is die verdeling redelik positief skeef.

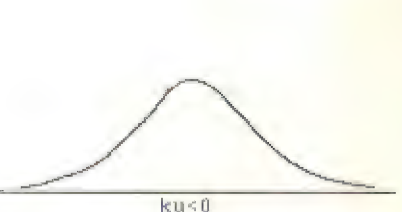
5.4.2 Kurtose (spitsheid)

'n Maat van die spitsheid van 'n verdeling is

$ku = \left[\frac{1}{2}\left(\frac{q_3 - q_1}{d_9 - d_1}\right) - 0,263\right]$

As ku > 0 is die verdeling spits en as ku < 0 is die verdeling plat.





Vir die voorbeeld is

$q_3 = 50,4$

$q_1 = 21,5$

$d_9 = 66,0$

$d_1 = 12,3$

en $ku = \frac{1}{2}\left(\frac{50,4 - 21,5}{66,0 - 12,3}\right) - 0,263$

$= 0,0061$

OPGAWE 5

1. Vir opgawes 4.1, 4.3, 4.8 en 4.9 bereken $\bar{x}$, Mo, Me, P_{30}, sk(P) en ku.

2. Die volgende tabel gee die punte (%) wat 400 studente in 'n Statistiektoets behaal het:

Punt	Aantal studente
[0 ; 10)	8
[10 ; 20)	19
[20 ; 30)	36
[30 ; 40)	73
[40 ; 50)	96
[50 ; 60)	62
[60 ; 70)	55
[70 ; 80)	25
[80 ; 90)	19
[90 ; 100)	7

i) Skets die frekwensieveelhoek.

ii) Bereken die

a) gemiddeldepunt
b) modale punte
c) die standaardafwyking
d) mate van skeefheid en kurtose.

3. Die volgende tabel gee die massas van 120 valskermsoldate.

Massa	Frekwensie
[65 ; 70)	8
[70 ; 75)	24
[75 ; 80)	32
[80 ; 85)	28
[85 ; 90)	20
[90 ; 95]	8

Bereken: i) $\bar{x}$ iv) Me
ii) S v) P_{25} en d_8
iii) V vi) sk(P) en ku

4. Bepaal die Mo, Me, $d_{[illegible]}$ en $P_{[illegible]}$ van opgawes 5.2 en 5.3 grafies.

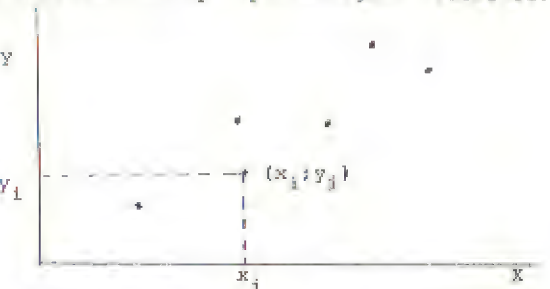
HOOFSTUK 6

LINEÊRE KORRELASIE EN REGRESSIE

6.1 Algemeen

In baie statistiese werk word daar twee veranderlikes x en y op elke element in die steekproef gemeet.

Sulke data kan ons op 'n puntediagram voorstel:



In die praktyk is dit so dat daar in baie gevalle een of ander verband tussen X en Y bestaan. In baie gevalle is die verband lineêr:

$$y = a+bx$$

Indien die verband lineêr is is die vraag wat is die vergelyking van die reguitlyn wat die verband die beste beskryf en hoe sterk is hierdie verband?

Die reguitlyn wat by sulke data aangepas word noem ons die regressielyn ($\hat{y}$ = a+bx) en die korrelasiekoëffisiënt (r) meet die sterkte van die verband.

6.2 Die Regressielyn

By 'n datastel $\{(x_i;y_i),\ i=1;\ \ldots;n\}$ is die kleinstekwadrate regressielyn

$$\hat{y} = a+bx$$

met

$$b = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} \qquad \text{(die helling van die lyn)}$$

en $a = \bar{y} - b\bar{x}$ (die afsnit op die y-as)

6.3 Die korrelasiekoëffisiënt (r)

'n Maatstaf van die sterkte van die verband tussen X en Y is

$$r = \frac{b\sqrt{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}}{\sqrt{\sum_{i=1}^{n} y_i^2 - n\bar{y}^2}}$$

r is altyd 'n getal wat tussen +1 en -1 lê. Hoe nader r aan +1 lê hoe 'n sterker positiewe lineêre verband bestaan daar tussen X en Y. Hoe nader r aan -1 lê hoe 'n sterker lineêre negatiewe verband bestaan tussen X en Y. As r naby 0 lê bestaan daar geen lineêre verband tussen X en Y nie.



Positiewe verband   Negatiewe verband   Geen verband

Voorbeeld 1

Die volgende data gee die verband tussen druk (x) en volume (y) in 'n sekere stelsel.

x	9	13	6	11	16
y	233	148	181	265	293

Pas 'n kleinstekwadrate regressielyn aan by die data en bereken die korrelasiekoëffisiënt, skets ook die lyn op 'n punte diagram.

Oplossing

BEREKENINGSTABEL

x	y	x²	y²	xy
9	233	81	54289	2097
13	148	169	21904	1924
6	181	36	32761	1086
11	265	121	70225	2915
16	293	256	85849	4688
$\Sigma x=55$	$\Sigma y_i=1120$	$\Sigma x_i^2=$	$\Sigma y_i^2=$	$\Sigma x_i y_i=$
$\bar{x}=11$	$\bar{y}=224$	663	265028	12710

$$b = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\Sigma x_i^2 - n\bar{x}^2}$$

$$= \frac{12710-5(11)(224)}{663-5(11)^2}$$

$$= 6,72$$

$$a = \bar{y} - b\bar{x}$$

$$= 224-6,72(11)$$

$$= 150,08$$

die verband is

$$\hat{y} = 150,08+6,72x$$

$$r = \frac{b\sqrt{\Sigma x_i^2 - n\bar{x}^2}}{\sqrt{\Sigma y_i^2 - n\bar{y}^2}}$$

$$= \frac{6,72\sqrt{663-5(11)^2}}{\sqrt{265028-5(224)^2}}$$

$$= 0,43$$

r is redelik klein dus is die verband nie baie sterk nie.

HOOFSTUK 7

WAARSKYNLIKHEIDSLEER

7.1 Algemeen

Waarskynlikheidsleer vorm 'n baie belangrike deel van Statistiek. Die doel van hierdie gedeelte van die kursus is egter net om spesifieke waarskynlikhede te bereken.

Die begrip waarskynlikheid kom in baie situasies voor maar dit sal beperk word tot waarskynlikhede wat maklik verstaan kan word. Om hierdie rede word daar slegs van dobbelspelle gebruik gemaak.

7.2 Basiese begrippe

7.2.1 Statistiese eksperimente

Indien 'n eksperiment n ewe moontlik uitkomste het noem ons dit 'n statistiese eksperiment.

Voorbeeld 1

1. Rol 'n dobbelsteen 1 keer. Daar is 6 ewe moontlike uitkomste.
2. Trek 'n kaart uit 'n goed geskommelde pak speelkaarte. Hier is daar 52 ewe moontlike uitkomste.

7.2.2. Die versameling van al die moontlike uitkomste van 'n statistiese eksperiment word die steekproefruimte (S) genoem. Die aantal elemente in S word aangedui met #S.

Voorbeeld 1 (vervolg)

1. $S = \{1;2;3;4;5;6\}$
   $\#S = 6$
2. $S = \{A_S;\ K_S;\ldots;2_S;\ldots\ 2_K\}$
   $\#S = 52$

7.2.3 Gebeurtenisse (EVENT)

Enige deelversameling van S word 'n gebeurtenis genoem.

Gebeurtenisse word voorgestel deur hoofletters bv. A,B,C,....

Die aantal element in gebeurtenis A word voorgestel deur #A.

Voorbeeld 1 (vervolg)

1. Laat A die gebeurtenis wees die uitkoms as 'n dobbelsteen een keer gerol word is 'n ewe getal.

   $A = \{2;4;6\}$
   $\#A = 3$

2. Laat B die gebeurtenis wees om 'n hart te kry as een kaart uit 'n pak speelkaarte getrek word.

   $B = \{A_H;\ K_H;\ \ldots;\ 2_H\}$
   $\#B = 13.$

7.3 Die klassieke definisie van Waarskynlikheid

Laat A 'n gebeurtenis relatief tot 'n steekproef ruimte S wees.

Die waarskynlikheid op A, P(A), word gedefinieer as

$$P(A) = \frac{\#A}{\#S}$$

Voorbeeld 1 (vervolg)

1. P (kry 'n ewe getal as 'n dobbelsteen een keer gerol word)
   $= \frac{\#A}{\#S} = \frac{3}{6} = \frac{1}{2}$
2. P (kry 'n hart as een kaart uit 'n pak speelkaarte getrek word)
   $= \frac{\#B}{\#S} = \frac{13}{52} = \frac{1}{4}$

Voorbeeld 2

'n Bak bevat 8 kaartjies met die nommers 1 tot 8. Een kaartjie word uit die bak getrek. Wat is die waarskynlikheid dat die nommer op hierdie kaartjie deelbaar sal wees deur 3?

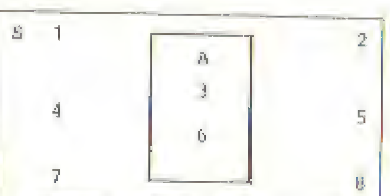
Oplossing:

$S = \{1;2;3;4;5;6;7;8\}$ $\quad \#S = 8$
A = nommer is deelbaar deur 3
$= \{3;6\}$ $\quad \#A = 2$
$P(A) = \frac{2}{8} = \frac{1}{4}$

7.4 Versamelingsleer

Aangesien gebeurtenisse deelversamelings van S is kan versamelingsleer direk toegepas word op waarskynlikhede. So byvoorbeeld kan Venn-diagramme gebruik word om gebeurtenisse voor te stel.

Voorbeeld 2 (vervolg)



7.5 Wette van waarskynlikheid

I. $0 \leqslant P(A) \leqslant 1$

As $P(A) = 0$ is A 'n onmoontlike gebeurtenis en as $P(A) = 1$ is A realiteit.

II. $P(A \cap B)$ = waarskynlikheid dat A en B sal plaasvind $= \frac{\#A \cap B}{\#S}$

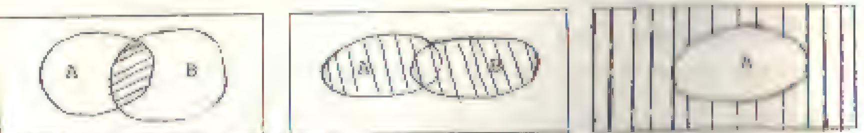
Indien $P(A \cap B) = 0$ is A en B onderling uitsluitend.
$P(A \cap B) = P(A)P(B)$ as en slegs as A en B onafhanklik is.

III. $P(A \cup B)$ = waarskynlikheid dat A of B sal plaasvind.
$= P(A) + P(B) - P(A \cap B)$.

IV. $P(\bar{A})$ = waarskynlikheid dat A nie sal plaasvind nie.
$= 1 - P(A)$

V. $P(A|B)$ = voorwaardelike waarskynlikheid op A gegee B het alreeds plaasgevind.
$= \frac{P(A \cap B)}{P(B)}$ mits $P(B) > 0$

Voorbeeld 3



Voorbeeld 4

Een kaart word uit 'n goed geskommelde pak speelkaarte getrek. Wat is die waarskynlikheid dat dit:

a) rooi
b) 'n vier
c) 'n rooi vier
d) rooi of 'n vier sal wees
e) As dit rooi is wat is die waarskynlikheid dat dit 'n vier sal wees.

Oplossing:

$S = \{A_S; \ldots; 2_K\}$ $\#S = 52$

A = kaart is rooi
$= \{A_H; \ldots; 2_H; A_D; \ldots; 2_D\}$ $\#A = 26$

B = kaart is 'n vier
$= \{4_S; 4_H; 4_D; 4_K\}$ $\#B = 4$

$A \cap B$ = kaart is rooi en 4
$= \{4_H; 4_D\}$ $\#A \cap B = 2$

a. $P(A) = \frac{\#A}{\#S} = \frac{26}{52} = \frac{1}{2}$

b. $P(B) = \frac{\#B}{\#S} = \frac{4}{52} = \frac{1}{13}$

c. $P(A \cap B) = \frac{\#A \cap B}{\#S} = \frac{2}{52} = \frac{1}{26}$

d. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
$= \frac{26}{52} + \frac{4}{52} - \frac{2}{52} = \frac{28}{52}$

e. $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{2}{52}}{\frac{26}{52}} = \frac{2}{26} = \frac{1}{13}$

7.6 Meerstadiumeksperimente

Indien 'n eksperiment meer as een stadium het is

$\#S = \#(\text{Eerste stadium}) \times \#(\text{2de stadium}) \times \ldots$

Voorbeeld 5

1. 'n Dobbelsteen word gerol en daarna word 'n muntstuk in die lug geskiet.

$\#S = 6 \times 2 = 12$

2. 'n Dobbelsteen word drie keer na mekaar gerol.

$\#S = 6 \times 6 \times 6 = 216$

7.7 Tel van elemente

I. Die aantal maniere waarop n element in 'n ry geplaas kan word is

$n!$ (n-fakulteit)

$= n \times (n-1) \times (n-2) \times \cdots \times 1$

II. Indien daar uit n elemente m elemente met terugplasing gekies word kan dit op n^m maniere gedoen word.

III. Indien daar uit n elemente m elemente sonder terugplasing gekies word en die volgorde waarin die elemente voorkom is belangrik kan dit op

$nPm = \frac{n!}{(n-m)!}$ maniere gedoen word.

IV. Indien daar uit n elemente m elemente sonder terugplasing gekies word en die volgorde is nie belangrik nie kan dit op

$nCm = \frac{n!}{m!(n-m)!}$ maniere gedoen word.

Voorbeeld 6

a. Die aantal maniere waarop 6 persone in ŉ ry kan sit is

$6! = 720.$

b. Die aantal maniere waarop 6 persone in ŉ ry kan sit, as drie van daardie persone langs mekaar moet sit is

$4! \times 3! = 144$

c. Die aantal maniere waarop ŉ rugbyspan van 8 voorspelers en 7 agterspelers uit 10 voorspelers en 9 agterspelers gekies kan word is

$10P8.9P7$

(Hier word aangeneem dat al die voor- en agterspelers in enige posisie kan speel.)

d. ŉ Spyskaart bestaan uit 3 voorgeregte, 6 hoofgeregte en 4 nageregte. Indien ŉ persoon 1 voorgereg, 3 hoofgeregte en twee nageregte mag kies het hy

$3C1 \times 6C3 \times 4C2 = 360$ keuses vir ŉ maaltyd.

Voorbeeld 7

ŉ Dobbelsteen word 2 keer na mekaar gerol.

Wat is die waarskynlikheid dat

a) die som sewe sal wees?
b) die som nie sewe sal wees nie?

Oplossing:

S = kies 2 uit 6 met terugplasing

$\#S = 6^2 = 36$

a. A = som is 7

$= \{(1;6);\ (6;1);\ (5;2);\ (2;5);(4;3);(3;4)\}$

$\#A = 6$

$P(A) = \frac{6}{36} = \frac{1}{6}$

b. $P(\bar{A}) = 1 - P(A)$

$= 1 - \frac{1}{6}$

$= \frac{5}{6}$

Voorbeeld 8

Uit 10 dames en 15 mans moet ŉ voorsitter en ondervoorsitter gekies word. Wat is die waarskynlikheid dat:

a) beide dames
b) die voorsitter ŉ man en die ondervoorsitter ŉ dame sal wees?
c) As die voorsitter ŉ man moet wees wat is die waarskynlikheid dat die ondervoorsitter ook ŉ man sal wees?

Oplossing S = kies 2 uit 25 sonder terugplasing volgorde belangrik.

a. $\#S = 25P2 = \frac{25!}{23!} = 600$

A = kies 2 uit 10

$\#A = 10\ P2 = \frac{10!}{8!} = 90$

$P(A) = \frac{90}{600} = \frac{3}{20}$

b. $\#S = 25P2 = 600$

B = kies 1 uit 15 en 1 uit 10

$\#B = (15P1) \times (10P1) = \frac{15!}{14!} \times \frac{10!}{9!} = 150$

$P(B) = \frac{150}{600} = \frac{1}{4}$

c. A Voorsitter is ŉ man $P(A) = 15P1/25P1 = \frac{15}{25}$

B Ondervoorsitter is ŉ man $P(B) = 15P1/25P1 = \frac{15}{25}$

A∩B Voorsitter is ŉ man en Ondervoorsitter is ŉ man

$P(A \cap B) = 15P2/25P2 = \frac{15.14}{25.24}$

$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{15.14}{25.24}}{\frac{15}{25}} = \frac{14}{24}$

Voorbeeld 9

NB

Twee kaarte word gelyktydig uit ŉ pak speelkaarte getrek.

Wat is die waarskynlikheid dat

a) Die een swart en die ander rooi sal wees?

b) Beide 9 sal wees?

Oplossing:

S = kies 2 uit 52 sonder terugplasing. Volgorde nie belangrik nie.

$\#S = 52C2 = \frac{52!}{2!50!} = \frac{52.51}{2.1} = 1326$

a) A = kies 1 uit 26 en 1 uit 26.

$\#A = 26C1 \times 26C1 = \frac{26!}{1!25!} \times \frac{26!}{1!25!} = 26 \times 26 = 676$

$P(A) = \frac{676}{1326} = \frac{338}{663}$

b) B = kies 2 uit 4.

$\#B = 4C2 = \frac{4!}{2!2!} = \frac{4.3}{2.1} = 6$

$P(B) = \frac{6}{1326} = \frac{1}{221}$

7.8 Bayes se Reël

Indien $B_1, B_2, \cdots B_k$ onderling uitsluitende gebeurtenisse relatief is tot steekproefruimte S sodat $S = B_1 + B_2 + \cdots + B_k$ en A is 'n ander gebeurtenis relatief tot S dan is

i. $P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \cdots + P(B_k)P(A|B_k)$

$= \sum_{i=1}^{k} P(B_i)P(A|B_i)$

ii. $P(B_\ell|A) = \frac{P(B_\ell)P(A|B_\ell)}{P(A)}$

Voorbeeld 10

Dit is bekend dat 50% van 'n sekere bevolking swart hare het, 30% het bruin hare en 20% het blonde hare. As 8% van die persone met swart hare, 12% van die persone met bruin hare en 20% van die persone met blonde hare blou oë het wat is die waarskynlikheid dat 'n persoon met blou oë blonde hare sal hê?

Oplossing

A = persoon het blou oë
B_1 = persoon het swart hare
B_2 = persoon het bruin hare
B_3 = persoon het blonde hare

$P(B_1) = 0,5$ $\quad P(B_2) = 0,3$ $\quad P(B_3) = 0,2$
$P(A|B_1) = 0,08$ $\quad P(A|B_2) = 0,12$ $\quad P(A|B_3) = 0,2$

$P(A) = \sum P(B_i)\,P(A|B_i)$

$= 0,5(0,08) + (0,3)(0,12) + (0,2)(0,2)$

$= 0,116$

$P(B_3|A) = \frac{P(B_3)P(A|B_3)}{P(A)}$

$= \frac{(0,2)(0,2)}{0,116}$

$= 0,3448$

Op hoeveel maniere ⇒ TELREËLS

OPGAWE 7

1. Hoeveel 4 letter woorde kan van die letters van die woord SKAKELBEAMPTE gemaak word as

   i. daar geen beperking is nie?
   ii. al 4 letters verskillend moet wees?

2. Op hoeveel maniere kan 8 motors in 'n ry geparkeer word as

   i. daar geen beperking is nie?
   ii. 'n sekere twee motors mekaar moet volg?

3. Op hoeveel maniere kan 'n waar/vals toets met 10 vrae beantwoord word.

4. Hoeveel motors kan in die Transvaal geregistreer word onder die nuwe registrasie-stelsel? (Klinkers word nie gebruik nie).

5. Hoeveel verskillende groepe van 5 kaarte kan uit 'n pak speelkaarte gehaal word? (Sonder terugplasing)

6. Hoeveel 5 syfer getalle kan gevorm word van die syfers 0,2,3,5,7 en 9 as 'n syfer nie meer as 1 keer gebruik mag word nie.

7. As vir 'n sekere steekproef is $P(A) = 0,4$, $P(\bar{B}) = 0,6$ en $P(A|B) = 0,3$ bereken $P(A \cap B)$.

8. Drie balle word gelyktydig getrek uit 'n bak met drie rooi, vier wit en vyf swart balle.

   Wat is die waarskynlikheid dat:

   i. almal swart
   ii. een wit, een rooi en een swart sal wees.

9. 'n Man skiet twee skote met 'n geweer na 'n teiken. Die waarskynlikheid dat hy die teiken met die eerste skoot sal tref is 0,3, met die tweede skoot 0,5 en met beide skote 0,25. As hy die teiken met die eerste skoot getref het wat is die waarskynlikheid dat hy dit met die tweede ook sal tref?

10. Twee balle word na mekaar uit 'n bak met 8 rooi, 15 wit en 6 blou balle getrek. Wat is die waarksynlikheid dat hoogstens 1 bal rooi sal wees as die trekking

    i. met terugplasing
    ii. sonder terugplasing is?

11. Indien $P(A) = 0,3$ en $P(B) = 0,5$ en A en B onderling uitsluitend is bereken

    i. $P(A \cup B)$
    ii. $P(A')$
    iii. $P(A' \cap B)$. (Skets die venndiagram!)

12. ŉ Kaartspel word met vyf kaarte gespeel. Wat is die waarskynlikheid om

i. vyf kaarte van dieselfde soort te kry?
ii. enige 2,3,4,5 en 6 te kry?

13. ŉ Raad bestaande uit ŉ voorsitter, ŉ ondersoorsitter en 2 lede moet uit 20 mans en 10 dames gekies word. (Eers word die voorsitter, dan die ondervoorsitter en dan die 2 lede gekies). Wat is die waarskynlikheid dat die voorsitter ŉ man, die ondervoorsitter ŉ dame en die twee lede mans sal wees?

14. By ŉ sekere hospitaal is daar 150 pasiënte waarvan 90 siekte A het, 70 siekte B en 10 het beide die siektes.

As ŉ pasient siekte B het wat is die waarskynlikheid dat hy ook siekte A sal hê?

15. Op ŉ boekrak is daar 12 Afrikaanse en 7 Engelse boeke. 4 Boeke word vanaf die rak gekies sonder terugplasing. Wat is die waarskynlikheid om hoogstens 1 Engelse boek te kry?

16. In ŉ loting vir twee pryse, word daar 100 kaartjies verkoop. ŉ Persoon koop 3 van die kaartjies.

Wat is die waarskynlikheid dat hy

a. Beide die pryse sal wen?
b. Een van die pryse sal wen?

as

I. die wenkaartjies gelyktydig getrek word
II. eers die eerste prys en dan die tweede prys getrek word.

17. 500 produkte, wat op 3 monteerbande vervaardig is, is ondersoek en die aantal produkte was in 3 katagorië verdeel:

	Band			
Toestand	1	2	3	
Aanvaarbaar	129	135	163	427
Kan herstel word	10	12	37	59
Afgeskryf	1	3	10	14
	140	150	210	500

Wat is die waarskynlikheid dat

i. ŉ produk afgeskyf is
ii. ŉ produk van produksieband 1 of 2 afkom
iii. ŉ produk van band 3 afkom wat aanvaarbaar is
iv. ŉ produk wat afgeskryf moet word van band 3 afkom?

18. Indien $P(A_1) = 0{,}6$ $P(A_2) = 0{,}2$ $P(A_3) = 0{,}2$
$P(H|A_1) = 0{,}04$ $P(H|A_2) = 0{,}01$ $P(H|A_3) = 0{,}03$

Bereken i. $P(H)$
ii. $P(A_1|H)$.

19. Die persentasies studente aan ŉ sekere Technikon wat Wiskunde, Statistiek en Rekenaarwetenskap neem is 30, 60 en 10% respektiewelik. Die waarskynlikheid dat ŉ student wat ŉ vak neem Engelsprekend is, is 0,12, 0,16 en 0,22 vir Wiskunde, Statistiek en Rekenaarwetenskap respektiewelik.

i. Wat is die waarskynlikheid dat ŉ student Engelsprekend is?
ii. Wat is die waarskynlikheid dat ŉ Engelssprekende student Wiskunde sal neem?

HOOFSTUK 8

WAARSKYNLIKHEIDSVERDELINGS

8.1 Verdelingsfunksie

Beskou 'n statistiese eksperiment met algemene uitkoms X. Die funksie wat waarskynlikhede toeken aan alle moontlike spesifieke uitkomste, x, word die verdelingsfunksie, $f(x) = P(X=x)$, genoem.

Vir diskrete veranderlikes bestaan f(x) uit 'n diskrete aantal punte (wat ons met 'n histogram benader) en vir kontinue veranderlikes is f(x) 'n gladde kromme.

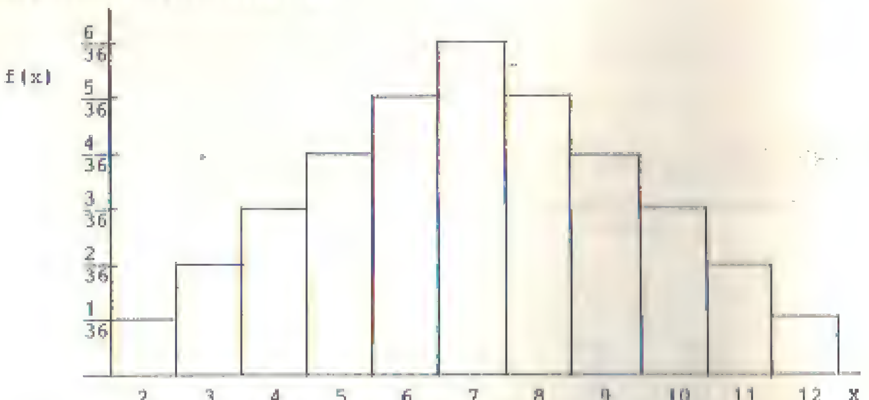
Voorbeeld 1

Twee onsydige dobbelstene word gelyk gerol en die som van die uitkomstes waargeneem.

Ons kan die resultate as volg opsom:

X=som	2	3	4	5	6	7	8	9	10	11	12
P(X=x)=f(x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Die verdelingsfunksie is:



Voorbeeld 2

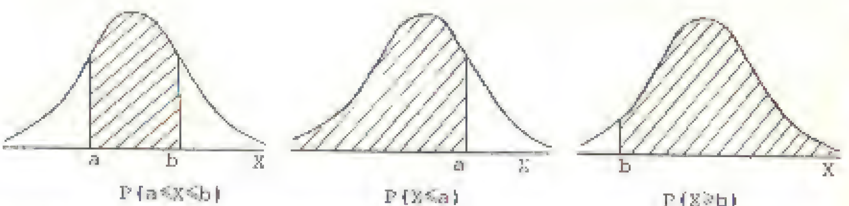
Die massas van 1000 persone word gemeet

'n Moontlike verdelingsfunksie is



8.1.1 Eienskappe van verdelingsfunksies

Die totale oppervlakte onder 'n verdelingsfunksie is presies een vierkante eenheid (sien voorbeeld 1). Om hierdie rede kan ons waarskynlikhede uitdruk as dele van die oppervlakte onder die verdelingsfunksie.



Sou ons dus die verdelingsfunksie ken kan ons baie maklik waarskynlikhede uit die verdeling bereken.

8.2 Kumulatieweverdelingsfunksie

Die kumulatieweverdelingsfunksie word gedefinieer as

$$F(x) = P(X \leq x).$$

Voorbeeld 3

Vir voorbeeld 1 is

X	2	3	4	5	6	7	8	9	10	11	12
f(x)=P(X=x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
F(x)=P(X≤x)	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	$\frac{36}{36}$

LET OP DAT - $P(X < x) = P(X \leq (x-1))$

$= 1 - P(X \geq x)$

$P(X \geq x) = 1 - P(X < x).$

8.3 Die Binomiaalverdeling

'n Statistiese eksperiment wat

i. uit n onafhanklike pogings bestaan
ii. die uitkoms slegs 'n sukses of mislukking is
iii. en die waarskynlikheid op 'n sukses in 'n enkel poging π is

word 'n binomiaal eksperiment genoem.

As X die aantal suksesse in so 'n eksperiment is, is 'n kort skryfwyse:

$X \sim b_1(n;\pi)$

Op die oog af lyk dit asof die binomiaalverdeling baie beperk is maar baie eksperimente kan tog wel as binomiaal geklassifiseer word.

Voorbeeld 4

i) ŉ Muntstuk word 20 keer gewerp en die aantal kere kruis word getel:

$X \sim b_i(20\ ;\ \frac{1}{2})$

ii) In ŉ bak is daar 7 wit en 3 swart balle. Vyf balle word na mekaar uit die bak met terugplasing getrek en die aantal swart balle word getel:

$X \sim b_i(5\ ;\ \frac{3}{10})$

iii) ŉ Produksie proses is so dat 10% van die produkte wat gemaak word defektief is. ŉ Steekproef van 25 produkte word geneem en die aantal defektiewes getel:

$X \sim b_i(25\ ;\ 0{,}10)$

8.3.1 Die verdelingsfunksie van die Binomiaalverdeling

As $X \sim b_i(n\ ;\ \pi)$

dan is $f(x) = P(X=x) = nCx\ \pi^x(1-\pi)^{n-x}$, $x = 0\ ;\ 1\ ;\ 2\ ;\ \ldots\ ;\ n$

Voorbeeld 5

ŉ Muntstuk word 20 keer gewerp. Wat is die waarskynlikheid om presies 5 keer kruis te kry?

$X \sim b_i(20\ ;\ \frac{1}{2})$

$P(X=5) = 20C5(\frac{1}{2})^5\ (1-\frac{1}{2})^{20-5}$

$= \frac{20!}{5!15!}\ (\frac{1}{2})^{20}$

$= 0{,}0148$

8.3.2 Die Kumulatiewe Binomiaalverdeling

As $X \sim b_i(n\ ;\ \pi)$

dan is $F(x) = P(X \leq x) = \sum_{r=0}^{x} nCr\ \pi^r(1-\pi)^{n-r}$

Voorbeeld 6

In ŉ bak is daar sewe wit en 3 swart balle. Vyf balle word na mekaar, met terugplasing, uit die bak getrek. Wat is die waarskynlikheid dat twee of minder van hierdie balle swart sal wees?

$X \sim b_i(5\ ;\ \frac{3}{10})$

$P(X \leq 2) = F(2) = P(X=0) + P(X=1) + P(X=2)$

$= 5C0(\frac{3}{10})^0(\frac{7}{10})^5 + 5C1(\frac{3}{10})^1(\frac{7}{10})^4 + 5C2(\frac{3}{10})^2(\frac{7}{10})^3$

$= 0{,}1681 + 0{,}3602 + 0{,}3087$

$= 0{,}8370$

In tabel X1 pp 16 van Stoker is die kumulatiewe binomiaalverdeling getabuleer vir sekere waardes van n en π.

Voorbeeld 7

$X \sim b_i(10\ ;\ 0{,}25)$

Uit tabel X1 van Stoker:

$P(X \leq 3) = F(3) = 0{,}7759$

$P(X < 5) = P(X \leq 4) = F(4) = 0{,}9219$

$P(X=6) = P(X \leq 6) - P(X \leq 5) = F(6) - F(5)$

$= 0{,}9965 - 0{,}9803$

$= 0{,}0162$

8.3.3 Eienskappe van die Binomiaalverdeling

As $X \sim b_i(n;\pi)$

dan is:

i. Die populasie gemiddeld, $\mu = n\pi$

ii. die populasie variansie, $\sigma^2 = n\pi(1-\pi)$.

Voorbeeld 8

10% van die produkte wat in ŉ sekere produksieproses gemaak word is defektief. ŉ Steekproef van 25 produkte word geneem.

i. Wat is die waarskynlikheid dat

a) persies 5
b) tussen 3 en 6 (albei ingesluit)

defektief sal wees?

ii. Wat is die gemiddeld en standaardafwyking van die aantal defektiewe produkte?

$X \sim b_1(25\ ;\ 0,10)$

i.a. $P(X=5) = F(5) - F(4)$

$= 0,9666 - 0,9020$

$= 0,0646$

b. $P(3 \leq X \leq 6) = P(X \leq 6) - P(X < 3)$

$= F(6) - F(2)$

$= 0,9905 - 0,5371$

$= 0,4534$

ii. $\mu = n\pi = 25(0,1) = 2,5$

$\sigma = \sqrt{n\pi(1-\pi)} = \sqrt{25(0,1)(0,9)}$

$= 1,5$

Voorbeeld 9

'n Steekproef van 800 families met 5 kinders word geneem. In hoeveel van die families verwag ons dat daar presies 3 seuns is?

$X \sim b_1(5\ ;\ \frac{1}{2})$

$P(X=3) = 5C3(\frac{1}{2})^3(\frac{1}{2})^2 = 0,3125$

Verwagtingswaarde = 800(0,3125) = 250.

8.4 Die Poissonverdeling (AVERAGE) 'n REEKS

As die aantal kere waargeneem word wat 'n sekere gebeurtenis plaasvind in 'n sekere interval of tydsinterval word die Poissonverdeling gebruik.

8.4.1 Die Verdelingsfunksie van die Poissonverdeling

Laat X die aantal suksesse wees in 'n sekere interval of tydsinterval en laat $\mu = \lambda = n\pi$ die gemiddeld van X wees. X besit 'n Poissonverdeling en in kort word dit geskryf as

$X \sim P_0(\lambda)$

Verder is $f(x) = P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$, $x = 0\ ;\ 1\ ;\ 2\ ;\ \ldots\ ;\ \infty$

∞ = tot in oneindigh

Voorbeeld 10 — X = kan tot in oneindig gaan.

Gemiddeld kom daar 11 defektiewe plekke voor in 'n rol materiaal wat 'n sekere masjien maak. Wat is die waarskynlikheid dat in 'n rol presies 5 defektiewe plekke sal voorkom?

Vir B_i moet jy weet HOEVEEL rolle is na gekyk, hier is slegs 'n aanduiding, dus is dit P_0

$X \sim P_0(11)$

$P(X=5) = \frac{e^{-11}11^5}{5!} = 0,0224$

8.4.2 Die Kumulatiewe Verdelingsfunksie van die Poissonverdeling

$$F(x) = P(X \leq x) = \sum_{r=0}^{x} \frac{e^{-\lambda}\lambda^r}{r!}$$

Hierdie funksie is getabuleer in tabel XII pp 19 van Stoker.

Voorbeeld 11

Gemiddeld betaal 9 kliënte per uur (interval) by 'n sekere kassier in 'n afdelingswinkel. Wat is die waarskynlikheid dat daar

a. 10 of minder
b. presies 7
c. meer as 9 kliënte per uur by die kassier sal betaal?

$X \sim P_0(9)$

Uit tabel XII:

a. $P(X \leq 10) = F(10) = 0,7060$

b. $P(X=7) = P(X \leq 7) - P(X \leq 6) = F(7) - F(6)$

$= 0,3239 - 0,2068$

$= 0,1171$

c. $P(X>9) = 1 - P(X \leq 9) = 1 - F(9)$

$= 1 - 0,5874$

$= 0,4126$

8.4.3 Benadering van die Binomiaal- deur die Poissonverdeling

Gewoonlik is dit makliker om Poisson- as Binomiaalwaarskynlikhede te bereken.

As $X \sim b_1(n;\pi)$ en n is redelik groot en π redelik klein kan ons met groot sukses waarskynlikhede benader met $X \sim P_0(\lambda)$, $\lambda = n\pi$.

Voorbeeld 12

Dit is bekend dat 1% van 'n besending mieliesaad nie sal ontkiem nie. Wat is die waarskynlikheid dat in 'n monster van 200 van hierdie mieliepitte minder as 3 sal ontkiem. Gebruik die Poissonbenadering.

$X \sim b_i(200\ ;\ 0{,}01)$

$\mu = \lambda = n\pi = 2$

$X \sim P_0(2)$

$P(X<3) = P(X=0) + P(X=1) + P(X=2)$

$= \frac{e^{-2}2^0}{0!} + \frac{e^{-2}2^1}{1!} + \frac{e^{-2}2^2}{2!}$

$= 0{,}1353 + 0{,}2707 + 0{,}2707$

$= 0{,}6767$ (Kan ook uit tabel XII van Stoker afgelees word.) (Die eksakte antwoord is 0,67667).

8.4.4 Eienskappe van die Poissonverdeling

As $X \sim P_0(\lambda)$

Dan is $\mu = \lambda$

en $\sigma^2 = \lambda$

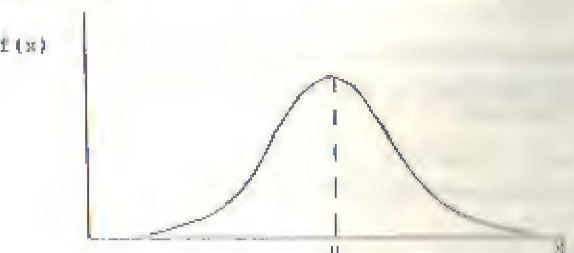
8.5 Die Normaalverdeling

Seker die mees belangrikste waarskynlikheidsverdeling is die Normaalverdeling.

NB In teenstelling met die vorige twee verdelings is die normaalverdeling 'n kontinue verdeling met die grootste verskil dat vir 'n kontinue verdeling geld $P(X\leq x) = P(X<x)$ of $P(X=x) = 0$. NB

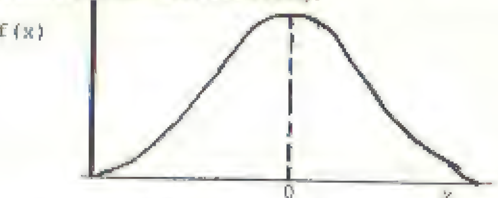
Die normaalverdeling het twee parameters wat bekend moet wees as ons waarskynlikhede wil bereken nl. μ en σ^2, en kortweg word geskryf $X \sim n(\mu;\sigma^2)$.

Die verdelingsfunksie van die normaalverdeling het 'n baie ingewikkelde vergelyking, wat ons nie sal bespreek nie, en lyk as volg:



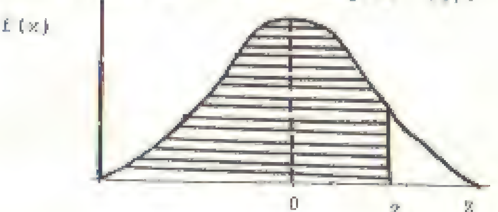
8.5.1 Die Standaardnormaalverdeling (Z)

'n Normaalverdeling waarvoor geld dat $\mu=0$ en $\sigma^2=1$ is die sogenaamde Standaardnormaalverdeling.



Kortweg word geskryf $Z \sim n(0;1)$.

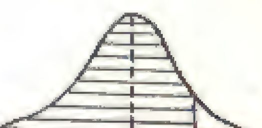
Oppervlaktes onder die standaardnormaalkrommes is getabuleer in tabel I van Stoker. Die tabel gee altyd die oppervlakte van $-\infty$ tot by z (waar z 'n positiewe getal is).



Voorbeeld 13

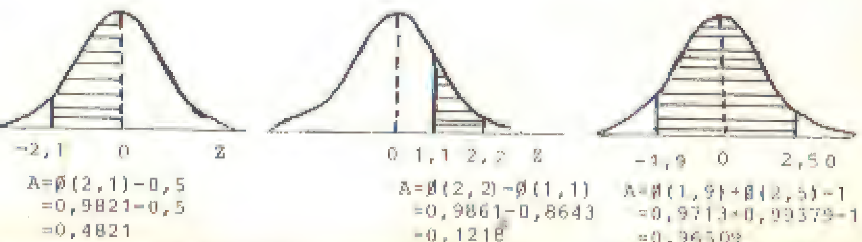


$A=\emptyset(1{,}96)$
$=0{,}9750$



$A=\emptyset(2{,}65)$
$=0{,}99598$

Omdat die totale oppervlakte onder hierdie kromme 1 vierkante eenheid is en omdat die kromme simetries is kan ons verskeie dele van die oppervlakte bereken.



$A=\emptyset(2{,}1)-0{,}5$
$=0{,}9821-0{,}5$
$=0{,}4821$

$A=\emptyset(2{,}2)-\emptyset(1{,}1)$
$=0{,}9861-0{,}8643$
$=0{,}1218$

$A=\emptyset(1{,}9)+\emptyset(2{,}5)-1$
$=0{,}9713+0{,}99379-1$
$=0{,}96509$

8.5.2 Standaardisasie

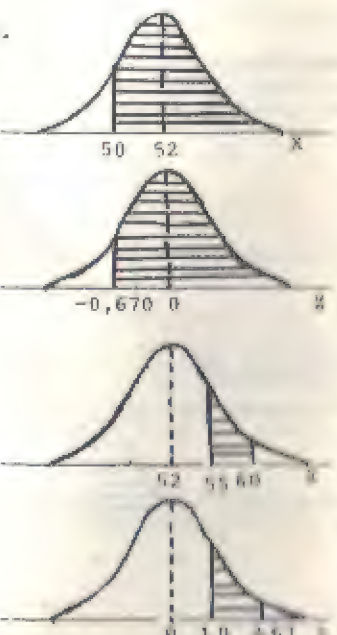
As $X \sim n(\mu; \sigma^2)$ en ons bereken $Z = \frac{X-\mu}{\sigma}$ dan is $Z \sim n(0;1)$.

Met ander woorde ons kan waarskynlikhede uit 'n normaalverdeling bereken deur net eers te standaardiseer en dan die oppervlakte, wat 'n seker waarskynlikheid is, uit tabel I af te lees.

Voorbeeld 14

Die punte in 'n sekere eksamen besit 'n normaalverdeling met gemiddeld $\mu=52$ en varainsie $\sigma^2=9$. Wat is die waarskynlikheid dat die punt van 'n sekere student

a. meer as 50
b. tussen 55 en 60 sal wees.



a. $X \sim n(52; 3^2)$

$p(X>50) = A$

$z = \frac{50-52}{3} = -0,67$

$p(X>50) = A = \emptyset(0,67)$
$=0,7486$

b. $p(55<X<60)$

$z_1 = \frac{55-52}{3}$ $\quad$ $z_2 = \frac{60-52}{3}$

$= 1,00$ $\quad$ $= 2,67$

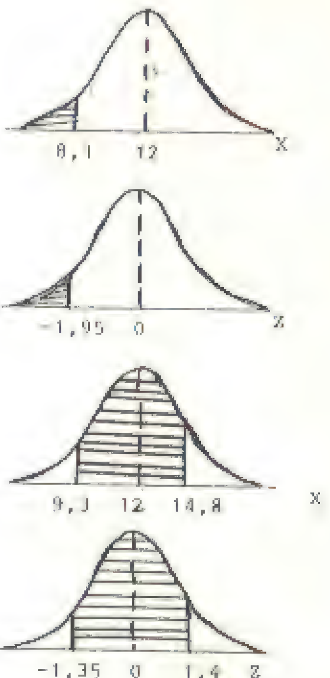
$p(55<X<60) = A$
$= \emptyset(2,67) - \emptyset(1,00)$
$= 0,99621 - 0,8413$
$= 0,15491$

Voorbeeld 15

Beskou $X \sim n(12\ ;\ 2^2)$

Bereken a. $P(X<8,1)$

b. $P(9,3<X<14,8)$



a. $P(X<8,1)$

$z = \frac{8,1-12}{2}$

$= -1,95$

$P(X<8,1) = 1-\emptyset(1,95)$
$= 1-0,9744$
$= 0,0256$

b. $P(9,3<X<14,8)$

$z_1 = \frac{9,3-12}{2}$ $\quad$ $z_2 = \frac{14,8-12}{2}$

$= -1,35$ $\quad$ $= 1,4$

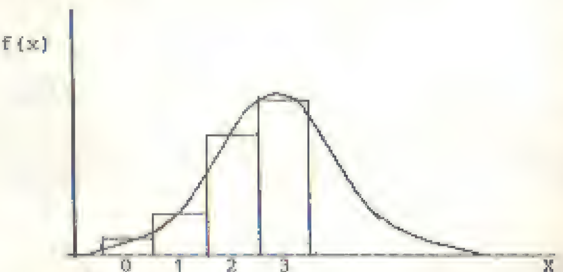
$P(9,3<X<14,8)$
$= \emptyset(1,35) + \emptyset(1,4) - 1$
$= 0,9115 + 0,9192 - 1$
$= 0,8307$

8.5.3 Benadering van die Binomiaal-deur die Normaalverdeling

As $X \sim b_1(n;\pi)$ en n is redelik groot kan ons waarskynlikhede benader uit $X \sim n(\mu;\sigma^2)$ met $\mu = n\pi$ en $\sigma^2 = n\pi(1-\pi)$.

Omdat ons van 'n diskrete verdeling oorgaan na 'n kontinue verdeling moet ons 'n kontinuïteitskorreksie maak deur 0,5 by te tel of af te trek voordat ons standaardiseer.

Voorbeeld 16



Gestel ons wil $P(X \leq 2)$ vir 'n sekere binomiaalverdeling bereken (sien histogram) deur die normaalbenadering. Die beste benadering sal gekry word as ons 2,5 (halfpad tussen 2 en 3) standaardiseer (sien gladde kromme) en dan die waarskynlikheid bereken.

Voorbeeld 17

Beskou $X \sim b_1(100 ; 0,15)$. Bereken:

a. $P(X \leq 20)$
b. $P(X \geq 23)$
c. $P(X = 13)$ deur 'n normaalbenadering:

$\mu = n\pi = 100(0,15) = 15$ $\sigma^2 = n\pi(1-\pi)$

$= 100(0,15)(0,85) = 12,75$

$X \sim n(15 ; 12,75)$

a. $z = \dfrac{(20 + 0,5) - 15}{3,57} =$

$= 1,54$



$P(X \leq 20) = \emptyset(1,54)$

$= 0,9382$

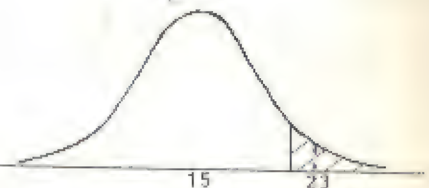
b. $z = \dfrac{(23 - 0,5) - 15}{3,57}$

$= 2,10$

$P(X \geq 23) = 1 - \emptyset(2,10)$

$= 1 - 0,9821$

$= 0,0179$

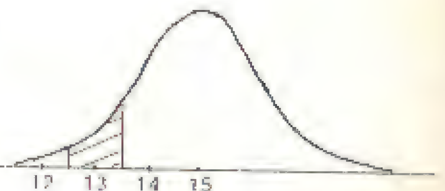


c. $z_1 = \dfrac{(13 - 0,5) - 15}{3,57}$

$= -0,70$

$z_2 = \dfrac{(13 + 0,5) - 15}{3,57}$

$= -0,42$



$P(X = 13) = \emptyset(0,7) - \emptyset(0,42)$

$= 0,7580 - 0,6628$

$= 0,0952$

(Hoekom is $P(X=13)$ nie nul nie?)

OPGAWE 8

1. Dit is bekend dat 3 uit elke 10 motorbestuurders veiligheidsgordels dra. Wat is die waarskynlikheid dat in 'n kettingbotsing van 8 motors

   a. geen een
   b. minder as 3
   c. meer as 3 bestuurders veiligheidsgordels aan sal hê?

2. Gemiddeld breek daar 4,1 voertuie uit 'n vloot van 15 voertuie per week. Wat is die waarskynlikheid dat in 'n sekere week

   a. geen
   b. meer as 3
   c. 1 of 2 voertuie sal breek?

3. Die breeksterkte van 'n sekere soort staalkabel besit 'n normaalverdeling wat gemiddeld $\mu = 102$ ton en 'n standaardafwyking van $\sigma = 15$ ton. As 'n myn die kabel sal gebruik wat is die waarskynlikheid dat dit sal breek onder 'n las van

   a. minder as 80 ton
   b. meer as 120 ton.

4. Dit is bekend dat 20% van die populasie van 'n sekere Afrika Staat aan 'n sekere siekte ly. Wat is die waarskynlikheid dat in 'n steekproef van 20 mense

   a. Presies 2
   b. Minder as 4
   c. Meer as 3 mense aan die siekte sal ly?

5. 'n Tikster maak gemiddeld 4 foute per bladsy wat sy tik. Wat is die waarskynlikheid dat sy, op 'n spesifieke bladsy

   a. Presies 4
   b. Minder as 9
   c. Meer as 6 foute sal maak.

6. Die punte wat 'n groep studente in 'n toets behaal het besit 'n normaalverdeling met gemiddeld $\mu = 55\%$ en variansie $\sigma^2 = 49\%$.

   Wat is die waarskynlikheid dat 'n student

   a. tussen 62% en 74%
   b. minder as 40
   c. meer as 80% sal behaal?

7. As $X \sim b_i(100;\ 0,31)$

   Bereken a. $P(X \geq 40)$
   b. $P(25 < X \leq 37)$

   met die normaalbenadering.

8. As $X \sim P_0(10)$

   Bereken a. $P(X = 5)$
   b. $P(7 \leq X \leq 11)$
   c. $P(X > 4)$

9. As $X \sim n(25,2;\ 3,7)$

   Bereken a. $P(X \geq 30)$
   b. $P(19 < X \leq 22)$.

10. Beskou:
   $X \sim n(14;16)$
   a. $P(X < 9)$
   b. $P(8 < X < 17)$
   c. $P(X > 7)$

HOOFSTUK 9

BERAMING

9.1.1 Beraming van die populasie gemiddeld, μ , by groot steekproewe (n>30)

Beskou ŉ populasie, met gemiddeld μ en variansie σ^2, met enige verdeling.

Vir n waarnemings, $X_1, X_2, \cdots, X_n$ uit die populasie is

$$\bar{X} = \frac{1}{n}\Sigma X_i.$$

Volgens die Sentralelimietstelling:

i. In die geval van ŉ normaalverdeelde populasie het $\bar{X}$ ŉ normaalverdeling met gemiddeld μ en variansie σ^2/n.

di. $\bar{X} \sim n(\mu ; \frac{\sigma^2}{n})$

ii. In die geval van ŉ nie-normaalverdeelde populasie het $\bar{X}$ ŉ normaalverdeling met gemiddeld μ en variansie σ^2/n, mits n groot is.

di. $\bar{X} \stackrel{\cdot}{\sim} n(\mu ; \frac{\sigma^2}{n})$

Die vraag is nou hoe goed beraam $\bar{x}$ vir μ. Omdat ons nooit μ eksak kan bereken nie is dit beter om ŉ sekere betroubaarheids interval te bereken (BI) waaraan ŉ waarskynlikheid gekoppel word. Verder word σ^2 deur s^2 vervang.

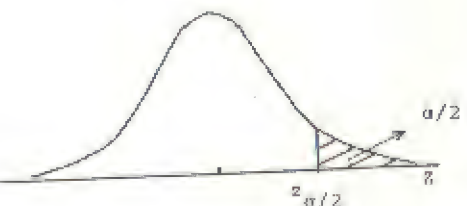
Die (1-α)% BI vir μ word gegee deur:

$$\left(\bar{x} - z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} ; \bar{x} + z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}\right)$$

waar

$\phi(z_{\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$. (afgelees uit tabel I of II van Stoker)

of $z_{\alpha/2} = \phi^{-1}(1 - \frac{\alpha}{2})$



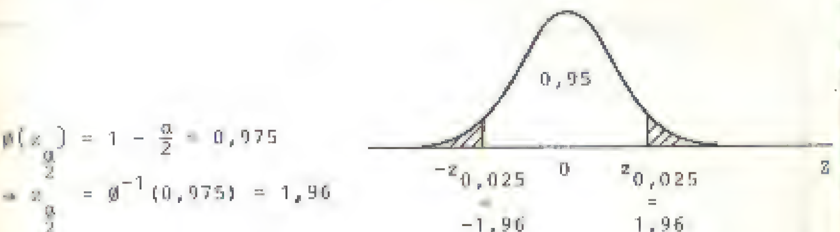
Voorbeeld 1

ŉ Steekproef van 100 elemente word uit ŉ normaalverdeelde populasie getrek en $\bar{x}$ = 15,2 en S^2 = 3,7 is bereken.

ŉ 95% BI vir μ is:

1-α = 0,95

⇒ α = 1-0,95 = 0,05 $\frac{\alpha}{2}$ = 0,025. $1 - \frac{\alpha}{2}$ = 0,975



$\phi(z_{\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2} = 0,975$

$\Rightarrow z_{\frac{\alpha}{2}} = \phi^{-1}(0,975) = 1,96$

Dus die 95% BI vir μ is:

$$\left(\bar{x} - z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} ; \bar{x} + z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}\right)$$

$$= \left(15,2 - 1,96 \frac{\sqrt{3,7}}{\sqrt{100}} ; 15,2 + 1,96 \frac{\sqrt{3,7}}{\sqrt{100}}\right)$$

$$= (15,2 - 0,38 ; 15,2 + 0,38)$$

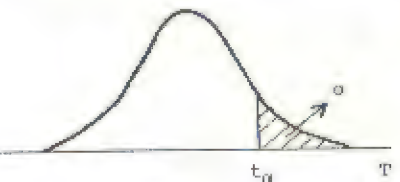
$$= (14,82 ; 15,58)$$

(Dit is 95% seker dat die werklike waarde van μ tussen 14,82 en 15,58 lê.)

9.1.2 Beraming van μ by klein steekproewe. n < 30

Hier kan die normaalverdeling nie gebruik word nie. Die Student-t verdeling, tabel III pp. 8 van Stoker lewer baie akkurater resultate.

Hierdie tabel gee t-waardes by betekenispeil α.



v = n-1 noem ons die vryheidsgrade (VG) van die verdeling.

Die (1 - α)% BI vir μ is

$$\left(\bar{x} - t_{\frac{\alpha}{2};v} \frac{S}{\sqrt{n}} ; \bar{x} + t_{\frac{\alpha}{2};v} \frac{S}{\sqrt{n}}\right)$$

waar $t_{\frac{\alpha}{2};v}$ ŉ tweekantig t-tabel waarde is by betekenispeil α, met

v = n-1 grade van vryheid.

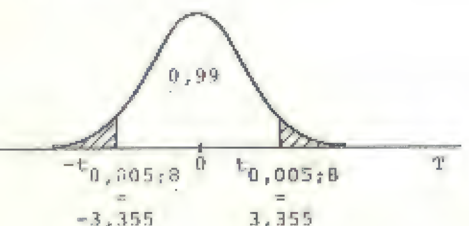
Voorbeeld 2

ŉ Steekproef van 9 blikkies groente is geneem en ŉ gemiddelde inhoud van 495 g met ŉ standaardafwyking van 1,27 g is waargeneem.

ŉ 99% BI vir μ is:

$$\alpha = 1 - 0,99 = 0,01 \quad \frac{\alpha}{2} = 0,005 \quad v = n-1 = 8$$

$$t_{\frac{\alpha}{2}} = 3,355$$



$$BI: \left(\bar{x} - t_{\frac{\alpha}{2};v}\frac{S}{\sqrt{n}} \; ; \; \bar{x} + t_{\frac{\alpha}{2};v}\frac{S}{\sqrt{n}}\right)$$

$$= \left(495 - 3,355\,\frac{1,27}{3} \; ; \; 495 + 3,355\,\frac{1,27}{3}\right)$$

$$= (495 - 1,42 \; ; \; 495 + 1,42)$$

$$= (493,58 \; ; \; 496,42)$$

Dit is dus byvoorbeeld 99% seker dat daar minder as 500 g in die blikkies is).

9.2 Beraming van populasie variansie, σ^2

$n(1-\alpha)$% BI vir die populasie variansie, σ^2, word gegee deur

$$\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2};v}} \; ; \; \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2};v}}\right)$$

waar $\chi^2_{\frac{\alpha}{2};v}$ en $\chi^2_{1-\frac{\alpha}{2};v}$ twee kantige chi-kwadraad waardes is

(afgelees uit tabel IV pp. 9 van Stoker) by $v = n-1$.

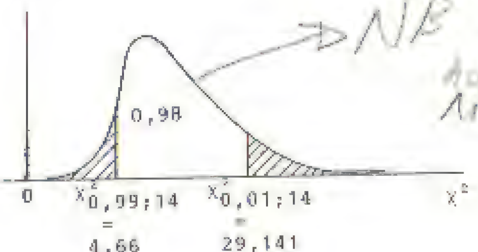
Voorbeeld 3

Die wanddruk van 15 staalpype is gemeet en $S^2 = 1,27$ is bereken.

ŉ 98% BI vir σ^2 is:

$$1-\alpha = 0,98 \quad \alpha = 0,02 \quad \frac{\alpha}{2} = 0,01 \quad 1 - \frac{\alpha}{2} = 0,99$$

$$v = n-1 = 14$$



$$BI: \left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2};v}} \; ; \; \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2};v}}\right)$$

$$= \left(\frac{14(1,27)}{29,141} \; ; \; \frac{14(1,27)}{4,66}\right)$$

$$= (0,610 \; ; \; 3,815)$$

9.3 Beraming van populasieverhoudings π (n altyd >30)

ŉ Volgende statistiek wat baie belangrik is is verhoudings. Hier neem ons ŉ steekproef van n elemente en tel die aantal elemente, X, wat tot ŉ sekere klas behoort.

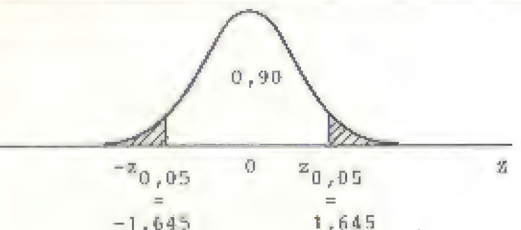
$P = \frac{X}{n}$ is die steekproef verhouding en π is die populasie-verhouding.

ŉ $(1-\alpha)$% BI vir π word gegee deur

$$\left(p - z_{\frac{\alpha}{2}}\sqrt{\frac{P(1-P)}{n}} \; ; \; p + z_{\frac{\alpha}{2}}\sqrt{\frac{P(1-P)}{n}}\right)$$

Voorbeeld 4

In 'n steekproef van 92 persone rook daar 15. 'n 90% BI vir die verhouding rokers is:



$$\left(\frac{15}{92} - 1{,}645\sqrt{\frac{\frac{15}{92}(1 - \frac{15}{92})}{92}}\ ;\ \frac{15}{92} + 1{,}645\sqrt{\frac{\frac{15}{92}(1 - \frac{15}{92})}{92}}\right)$$

$$= \left(\frac{15}{92} - 0{,}063\ ;\ \frac{15}{92} + 0{,}063\right)$$

$$= (0{,}100\ ;\ 0{,}226)$$

(Ons is 90% seker dat daar tussen 10% en 22,6% van die mense in die populasie rook.)

[illegible] 3

1. In 'n steekproef van 12 elemente was gevind dat die gemiddeld 11,[illegible] en die variansie 5,2 is.

   Gee 90 en 95% betroubaarheidintervalle vir die populasie gemiddeld μ en populasie variansie σ^2.

2. In 'n ondersoek van 140 artikels was daar 21 defek. Gee 95 en 99% betroubaarheidsintervalle vir die verhouding defektiewe artikels.

3. Beskou die volgende data:

   [illegible] [illegible] 3,3 2,9 3,2 3,5 3,0 2,1 1,9

   a. Beraam μ en σ^2.
   b. Gee 95% betroubaarheidsintervalle vir μ en σ^2.

4. Die Wiskunde- en Statistiekpunte vir 'n sekere groep studente in 'n toets was as volg:

Wiskunde :	46	63	77	86	84	93	79
Statistiek:	72	68	92	87	88	87	97

   Is daar 'n oorvleuling van die 95% betroubaarheidsintervalle van μ by die wiskunde en statistiekpunte?

5. Uit vraelyste wat aan 300 persone gestuur was is gevind dat 120 meer as R25 000,00 per jaar verdien. Gee 'n 95% BI. vir die verhouding persone wat minder as R25 000,00 verdien.

6. Die gemiddelde opbrengs van 70 mielieplante was 973,2 gr met 'n standaardafwyking van 64,3 gr. Gee 90 en 99% BI. vir μ.

HIPOTESETOETSING wo. 28 Jul 1993

10.1 Algemeen

Def {Die metode van statistiesebesluitneming word hipotesetoetsing genoem.

'n Hipotese is 'n sekere aanname wat gemaak word en dan word hierdie aanname getoets.

STAPPE:

1. 'n Steekproef word uit 'n populasie getrek en steekproefparameters soos $\bar{x}$, S^2 of P word bereken. 2. Nou word 'n aanname oor die betrokke populasieparameter, μ, σ^2 of π gemaak.
3. 'n Toetsstatistiek word bereken en teen 'n kritieke waarde, wat uit een van die verdelingstabelle in die boek van D.J. Stoker afgelees word, getoets. 4. Nou word die aanname wat gemaak is aanvaar of verwerp.

Steekproefparameters kan nooit presies dieselfde as populasieparameters wees nie. Daar is altyd 'n klein verskil as gevolg van kanseffekte en daar moet by hipotesetoetsing vir hierdie afwyking voorsiening gemaak word. Dit word met 'n waarskynlikheid gedoen en hierdie waarskynlikheid word die betekenispeil, α, genoem. Gewoonlik word α gekies as 0,1, 0,05, 0,01 of 0,001. (10%, 5%, 1% en 0,1%).

Die toetsstatistiek word met behulp van 'n sekere formule uit die beskikbare data bereken.

Daar word altyd twee hipoteses H_0 en H_1 gestel. As byvoorbeeld getoets wil word of 'n sekere populasiegemiddeld nogsteeds 15 is, is

Nul Hipotese H_0: $\mu = 15$

Alternatief H_1: $\mu \neq 15 \Rightarrow$ of $\mu > 12$ of $\mu < 12$

H_0 het altyd 'n = teken en H_1 een van die volgende tekens: $\neq$, $<$ of $>$. As H_1 'n $\neq$ teken het word die toets tweekantig, as dit 'n $<$ teken het lingseenkantig en as dit 'n $>$ teken het regseenkantig genoem. Uit die probleem wat opgelos moet word kan daar baie maklik besluit word watter teken by H_1 gebruik moet word.

Die uiteindelike beslissing van 'n hipotesetoets is of om H_0 te aanvaar of om dit te verwerp. As by bogenoemde voorbeeld H_0 aanvaar word is die populasiegemiddeld nog steeds 15 en as H_0 verwerp word is die populasiegemiddeld nie meer gelyk aan 15 nie.

Die foute wat by hipotesetoetsing gemaak kan word kan as volg voorgestel word:

[Afwyking van sp parameter word met wsk.heid α aangedui deur α - die betekenispeil.

		Werklike situasie	
		H_0 waar	H_0 nie waar nie
Uitslag van die toets:	Aanvaar H_0	Regte besluit	Tipe II-fout
	Verwerp H_0	Tipe I-fout	Regte besluit

Die waarskynlikheid op 'n tipe I-fout is gelyk aan die betekenispeil α van die toets en dit kan beheer word deur ander waardes aan α te gee. ← Hipotese word WAAR is
wsk.heid om tipe-I fout te maak - α

Om die waarskynlikheid op 'n tipe II-fout te bereken is baie moeilik maar as die steekproewe wat geneem word redelik groot is en die meting wat gemaak word baie akkuraat is kan aanvaar word dat hierdie waarskynlikheid klein sal wees.

Elke toetsstatistiek besit 'n sekere verdeling (Z, t, χ^2 of F). ("Chi-kwadraat") Met behulp van die betekenispeil word die verdelingskromme, van die betrokke verdeling, in twee gebiede, 'n aanvaardings- en verwerpingsgebied verdeel:

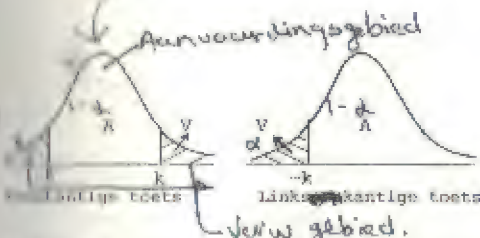
T_s = Toetsstatistiek

$-k < T_s < k$ ∴ Aanvaar H_0

$T_s \leq -k$ ∴ Verwerp H_0

of $T_s \geq k$ " "

Aanvaardingsgebied



Tweekantige toets — Linkseenkantige toets

Verw gebied.



Regseenkantige toets

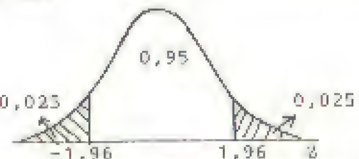
1 - α = aanv gebied
α - verw gebied
k - krit waarde

Die verwerpingsgebied word altyd so gekies dat die oppervlakte daarvan gelyk is aan die betekenispeil, α, van die toets. Die waarde(s) k, op die sketse, is kritiekewaardes en word uit die tabelle van die betrokke verdeling afgelees.

Dan word die toetsstatistiek met die kritiekewaarde(s) vergelyk. As die toetsstatistiek in die verwerpingsgebied val word H_0 verwerp en as dit in die aanvaardingsgebied val word H_0 aanvaar.

Voorbeeld 1

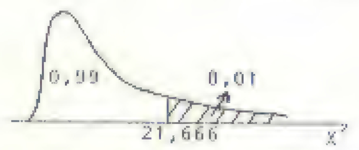
a) 'n Tweekantige kritieke waarde uit die z-verdeling by betekenispeil $\alpha = 0{,}05$ is:



As die toetsstatistiek tussen -1,96 en 1,96 lê word H_0 aanvaar anders word dit verwerp.

b. 'n Regseenkantige kritieke waarde uit die χ^2-verdeling met 9 vryheidsgrade, by betekenispeil $\alpha = 0{,}01$ is:

Toets vir Normaliteit



10.2 Toets vir Normaliteit

In baie van die toetse wat volg word die aanname gemaak dat die steekproewe uit normaalverdeelde populasies kom. Die vraag is nou hoe kan getoets word of die populasie wel normaal verdeel is.

As die steekproef redelik groot is kan 'n histogram geteken word om te kyk of dit ongeveer soos 'n normaalverdelingskromme lyk.

As die steekproef nie baie groot is nie moet die volgende toets gedoen word:

a. Rangskik die data van klein na groot.

b. Stip elke waarneming uit teenoor $(100i)/(n+1)$, waar i die nommer van die gerangskikte waarneming is, op normaalgrafiekpapier.

c. As die punte op die grafiek ongeveer op 'n reguitlyn lê kan ons aanneem dat die steekproef uit 'n normaalverdeelde populasie kom. Waar die punte baie afwyk van 'n reguitlyn of waar daar sistematiese afwykings is, is die populasie nie normaal verdeel nie. $\frac{100i}{n+1}$

Voorbeeld 2

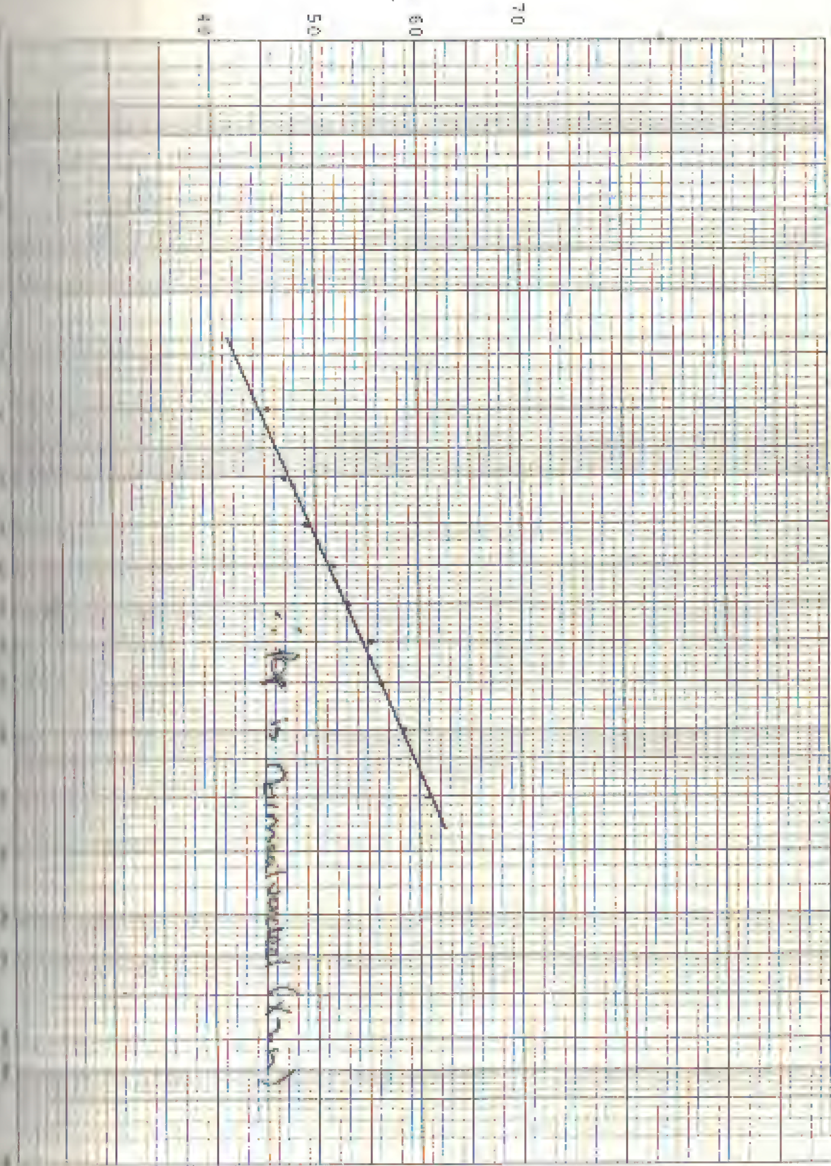
Gestel 'n steekproef van grootte 9 is geneem en die data was as volg:

49,0 53,1 46,9 56,1 45,1 60,5 54,9 51,8 58,5

Die volgende tabel word opgestel:

Data gerangskik (X)	45,1	46,9	49,0	51,8	53,1	54,9	56,1	58,5	60,5
i (nommer)	1	2	3	4	5	6	7	8	9
(100i)/(n+1)	10	20	30	40	50	60	70	80	90

$\frac{100i}{n+1}$



Nou stip ons X uit teenoor (100i)/(n+1) op normaalgrafiekpapier (sien grafiek 1)

Die punte lê baie naby aan 'n reguitlyn en daar kan aanvaar word dat die data uit 'n normaalverdeelde populasie kom.

10.3 Toets vir een Gemiddeld

Dit gebeur dikwels dat die populasie gemiddeld, μ, van 'n normaalverdeling bekend is. Nou word 'n steekproef uit die populasie geneem en $\bar{x}$ en S^2 word bereken. Die vraag is nou of die populasie gemiddeld nog steeds waarde $\mu = \mu_0$ het.

Hipoteses: $H_0 : \mu = \mu_0$

$H_1 : \mu \neq \mu_0$ of $\mu < \mu_0$ of $\mu > \mu_0$

Toetsgrootheid: $z_S = \frac{(\bar{x} - \mu_0)}{S}\sqrt{n}$ of $T_S = \frac{(\bar{x} - \mu_0)}{S}\sqrt{n}$

Kritieke waarde: $z_S \sim n(0;1)$ as $n \geq 30$ (tabel II van Stoker)

$T_S \sim t(n-1)$ as $n < 30$ (tabel III van Stoker)

as H_0 waar is.

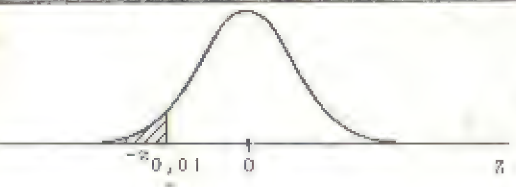
Voorbeeld 3

Die gemiddelde inhoud van 'n sekere soort ingelegde groente moet 500 g per blikkie wees. In 'n ewekansige steekproef van 40 blikkies is 'n gemiddelde inhoud van 496 g met 'n standaardafwyking van 20 g gevind. Toets by 'n 1% betekenispeil of daar nog genoeg groente in die blikkies is. Aanvaar dat die steekproef uit 'n normaalverdeelde populasie kom:

Toets vir 5% betekenispeil

$H_0 : \mu = 500$
$H_1 : \mu < 500$

$z_S = \frac{(\bar{x} - \mu_0)}{s}\sqrt{n} \sim n(0;1)$ onder H_0.



MAW: Aanvaar H_0 indien $z_s > -2,326$
Verwerp H_0 indien $z_s \leq -2,326$

$z_s = \frac{(496 - 500)}{20}\sqrt{40} = -1,26$

DWS: H_0 word aanvaar.
Daar is nog genoeg groente in die blikkies.

Voorbeeld 4

Doen dieselfde metode v voorb 2, blad.

Die gemiddelde leeftyd van 'n sekere soort battery behoort 12,5 uur te wees. In 'n ondersoek van 10 van die batterye is die volgende leeftye waargeneem:

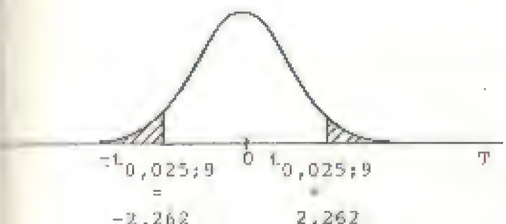
10,5 11,2 12,9 12,7 10,3 10,4 10,9 12,3 11,6 12,7

Toets by $\alpha = 0,05$ of daar 'n betekenvolle verskuiwing van die gemiddelde leeftyd is. (Neem aan dat die populasie normaalverdeel is.)

$H_0 : \mu = 12,5$
$H_1 : \mu \neq 12,5$

$T_S = \frac{(\bar{x} - \mu_0)}{s}\sqrt{n} \sim t$ onder H_0.

$[\bar{x} = 11,55 \quad s^2 = 1,06 \quad n = 10 \quad v = 9]$



MAW: Aanvaar H_0 indien $-2,262 < t_s < 2,262$
Verwerp H_0 indien $t_s \leq -2,262$ of $t_s \geq 2,262$

$t_s = \frac{(11,55 - 12,5)}{\sqrt{1,06}}\sqrt{10} = -2,92$

DWS: H_0 word verwerp.
Die leeftyd is nie meer 12,5 uur nie (Let op dat die uitspraak nie is of die leeftyd korter of langer as 12,5 uur is nie).

10.4 Toets vir een Verhouding

Die vraag is of 'n steekproefverhouding, $P = \frac{X}{n}$, nog steeds uit 'n normaalverdeelde populasie kom met populasieverhouding $\pi = \pi_0$ (X is die aantal elemente in die steekproef wat tot 'n sekere klas behoort).

Hipoteses: $H_0 : \pi = \pi_0$

$H_1 : \pi \neq \pi_0$ of $\pi < \pi_0$ of $\pi > \pi_0$

Toetsgrootheid: $z_S = \frac{P - \pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}$

Kritieke waarde: $z_S \sim n(0;1)$ $n \geq 30$ (tabel II van Stoker)

as H_0 waar is

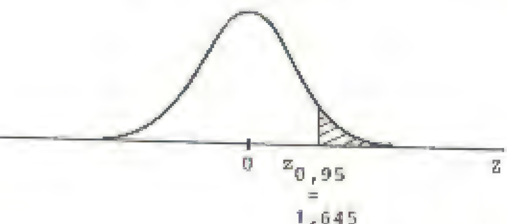
Voorbeeld 5

'n Vervaardiger beweer dat nie meer as 7% van die produkte wat hy maak defektief is nie. In 'n ondersoek van 125 van hierdie produk was daar 14 defektief. Kan die vervaardiger se eis ondersteun word by $\alpha = 0,05$? (Neem aan dat die populasie normaalverdeel is.)

$H_0 : \pi = 0,07$

$H_1 : \pi > 0,07$

$$z_S = \frac{P - \pi_0}{\sqrt{\pi_0(1-\pi_0)/n}} \sim n(0;1) \text{ onder } H_0.$$



MAW: Aanvaar H_0 indien $z_S < 1,645$.
Verwerp H_0 indien $z_S \geq 1,645$.

$$z_S = \frac{\frac{14}{125} - 0,07}{\sqrt{(0,07)(0,93)/125}} = 1,8404$$

DUS: H_0 word verwerp.
Daar is meer as 7% defektiewe produkte.

10.5 Toets vir een Variansie

Hier is die vraag of 'n steekproef, met variansie S^2, uit 'n normaalverdeelde populasie, met $\sigma^2 = \sigma_0^2$, kom.

Hipoteses: $H_0 : \sigma^2 = \sigma_0^2$

$H_1 : \sigma^2 \neq \sigma_0^2$ of $\sigma^2 < \sigma_0^2$ of $\sigma^2 > \sigma_0^2$

Toetsgrootheid: $X_S^2 = \frac{(n-1)S^2}{\sigma_0^2}$

Kritieke waarde: $X_S^2 \sim X^2(n-1)$ (tabel IV van Stoker)

as H_0 waar is.

Voorbeeld 6

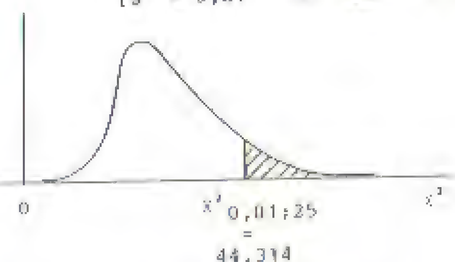
'n Produksieproses is ingestel om plastiekpype te vervaardig waarvan die variansie van die wanddikte nie meer as 0,15 mm² is nie. In 'n ondersoek van 26 monsters van die pype was die variansie van die wanddikte 0,21 mm². Toets by $\alpha = 0,01$ of die pype nog steeds aanvaarbaar is.

$H_0 : \sigma^2 = 0,15$

$H_1 : \sigma^2 > 0,15$

$$X_S^2 = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2 \text{ onder } H_0.$$

$[s^2 = 0,21 \quad n = 26 \quad v = 25]$



MAW: Aanvaar H_0 indien $X_S^2 < 44,314$
Verwerp H_0 indien $X_S^2 \geq 44,314$

$$X_S^2 = \frac{25(0,21)}{0,15} = 35,0$$

DUS: H_0 word aanvaar.
Die pype is nog aanvaarbaar.

10.6 Toets vir twee Gemiddeldes

10.6.1 Die doel hier is om te toets of twee onafhanklike steekproewe, met gemiddeldes $\bar{x}_1$ en $\bar{x}_2$, kom uit twee normaalverdeelde populasies, met gelyke variansies, waar $\mu_1 = \mu_2$.

Hipoteses: $H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$ of $\mu_1 < \mu_2$ of $\mu_1 > \mu_2$

Toetsgrootheid: $T_S = \dfrac{\bar{x}_1 - \bar{x}_2}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

met

$$S = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}}$$

Kritieke waarde: $T_S \sim t(n_1+n_2-2)$ $n_1<30$ $n_2<30$ (tabel III van Stoker)

as H_0 waar is.

Voorbeeld 7

Om te toets of daar 'n verskil tussen die gemiddelde slagmassas van twee hoenderrasse is, is twee onafhanklike steekproewe geneem en die resultate (slagmassas in kg) was as volg:

RAS A:	1,24	1,19	1,25	1,42	1,14	1,32	1,23	1,15	1,21
RAS B:	1,29	1,48	1,16	1,36	1,48	1,34	1,20	1,31	

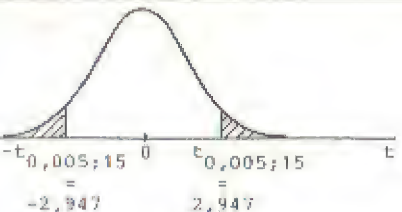
Aanvaar dat die steekproewe uit normaalverdeelde populasies kom met dieselfde variansie en toets of daar 'n verskil is tussen die gemiddelde slagmassas by $\alpha = 0{,}01$.

$H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$

$$T_S = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t \text{ onder } H_0.$$

$$\text{met } s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$[\bar{x}_1 = 1{,}24$ $\bar{x}_2 = 1{,}33$ $s_1^2 = 0{,}0076$ $s_2^2 = 0{,}0134$
$n_1 = 9$ $n_2 = 8$ $v = 15]$



HAW: Aanvaar H_0 indien $-2{,}947 < t_s < 2{,}947$

Verwerp H_0 indien $t_s \leq -2{,}947$ $t_s \geq 2{,}947$

$$s = \sqrt{\frac{8(0{,}0076) + 7(0{,}0134)}{9 + 8 - 2}} = 0{,}1015$$

$$t_s = \frac{1{,}24 - 1{,}33}{0{,}1015\sqrt{\frac{1}{9} + \frac{1}{8}}} = -1{,}825$$

GGD: H_0 word aanvaar.
Die twee steekproewe kom uit populasies met dieselfde gemiddeld

10.6.2 Hier wil getoets word of twee onafhanklike steekproewe, met gemiddeldes $\bar{x}_1$ en $\bar{x}_2$, kan kom uit normaalverdeelde populasies waar $\mu_1 = \mu_2$.

Hipoteses: $H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$ of $\mu_1 < \mu_2$ of $\mu_1 > \mu_2$

Toetsgrootheid: $Z_S = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$

Kritieke waarde: $Z \sim n(0;1)$ $n_1 \geq 30$ en $n_2 \geq 30$ (tabel II van Stoker) as H_0 waar is.

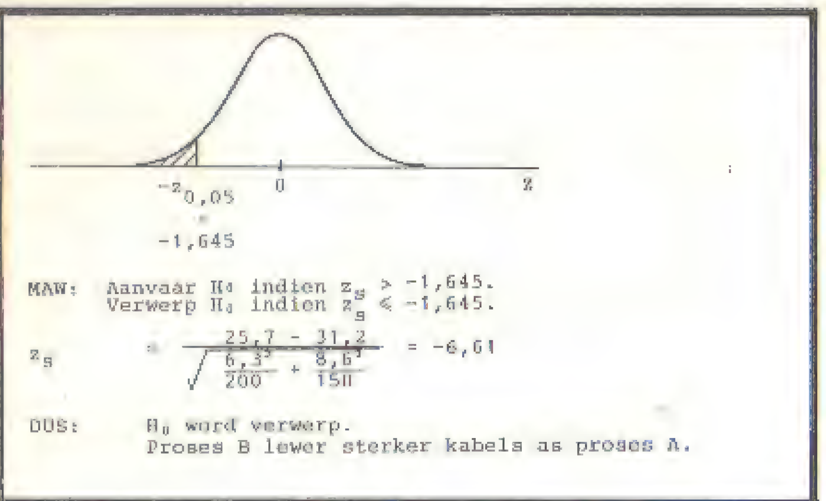
Voorbeeld 8

Twee prosesse, A en B, word gebruik om staalkabels te maak.

In 'n monster van 200 kabels van proses A is 'n gemiddelde breeksterkte van 25,7 ton met 'n standaardafwyking van 6,3 ton waargeneem en in 'n monster van 150 kabels van proses B is 'n gemiddelde breeksterkte van 31,2 ton met 'n standaardafwyking van 8,6 ton waargeneem. Toets by 'n 5% betekenispeil of proses B sterker kabels lewer as proses A. (Neem aan dat die steekproewe onafhanklik is en dat die populasies normaalverdeel is.)

$H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$

$$Z_S = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim n(0;1) \text{ onder } H_0.$$



MAW: Aanvaar H_0 indien $z_s > -1,645$.
Verwerp H_0 indien $z_s \leqslant -1,645$.

$$z_s = \frac{25,7 - 31,2}{\sqrt{\frac{6,3^2}{200} + \frac{8,6^2}{150}}} = -6,61$$

DUS: H_0 word verwerp.
Proses B lewer sterker kabels as proses A.

10.7 Toets vir twee verhoudings

Hier wil getoets word of twee steekproefverhoudings, P_1 en P_2, uit onafhanklike steekproewe, kom uit normaalverdeelde populasies waar $\pi_1 = \pi_2$.

Hipoteses: $H_0 : \pi_1 = \pi_2$

$H_1 : \pi_1 \neq \pi_2$ of $\pi_1 < \pi_2$ of $\pi_1 > \pi_2$

Toetsgrootheid: $$z_s = \frac{P_1 - P_2}{\sqrt{P(1-P)(\frac{1}{n_1} + \frac{1}{n_2})}}$$

waar

$$P = \frac{X_1 + X_2}{n_1 + n_2}$$

Kritieke waarde: $z_s \sim n(0;1)$ $n_1 \geqslant 30$ $n_2 \geqslant 30$ (tabel II van Stoker)

as H_0 waar is.

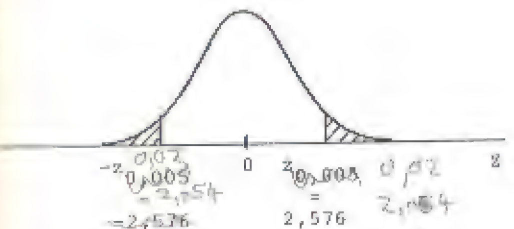
Voorbeeld 9

'n Opname het getoon dat uit 120 persone in stad A is daar 15 werkloos en uit 170 persone uit stad B is daar 16 werkloos. Toets by $\alpha = 0,01$ of daar 'n verskil is vir die werkloosheidsyfers van die twee stede.

$H_0 : \pi_1 = \pi_2$
$H_1 : \pi_1 \neq \pi_2$

$$z_s = \frac{P_1 - P_2}{\sqrt{P(1 - P)(\frac{1}{n_1} + \frac{1}{n_2})}} \sim n(0;1) \text{ onder } H_0.$$

met $P = \frac{X_1 + X_2}{n_1 + n_2}$



MAW: Aanvaar H_0 indien $-2,576 < z_s < 2,576$.
Verwerp H_0 indien $z_s \leqslant -2,576$ of $z_s \geqslant 2,576$

$$P = \frac{15 + 16}{120 + 170} = 0,107$$

$$z_s = \frac{\frac{15}{120} - \frac{16}{170}}{\sqrt{(0,107)(0,893)(\frac{1}{120} + \frac{1}{170})}} = 0,838$$

DUS: H_0 word aanvaar.
Die twee verhoudings is nie dieselfde nie.

10.8 Toets vir twee variansies

Hier wil getoets word of die variansies van twee onafhanklike steekproewe, uit normaalpopulasies, dieselfde is.

Hipoteses: $H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$ of $\sigma_1^2 < \sigma_2^2$ of $\sigma_1^2 > \sigma_2^2$

Toetsgrootheid: $F_S = \frac{S_1^2}{S_2^2}$

Kritieke waarde: $F_S \sim F\ (v_1 = n_1 - 1;\ v_2 = n_2 - 1)$ (tabelle V, VI VII van Stoker)

as H_0 waar is

(Hierdie tabelle gee regseenkantige waardes. Linkseenkantige waardes word verkry deur die rolle vir v_1 en v_2 om te ruil, en dan die omgekeerde van hierdie waarde te bereken. 'n Linkseenkantige 1% F-waarde by $v_1 = 10$ en $v_2 = 12$ is $\frac{1}{4,71} = 0,212$).

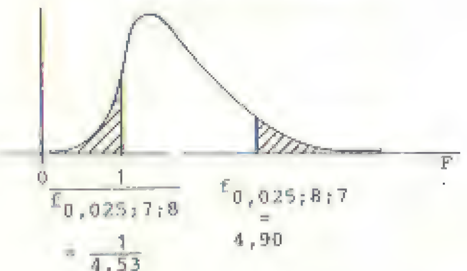
Voorbeeld 10

Toets of die variansies in voorbeeld 7 wel dieselfde is by $\alpha = 0,05$.

$H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$

$F_S = \frac{S^2}{S^2} \sim F$ onder H_0.



$\frac{1}{f_{0,025;7;8}} = \frac{1}{4,53} = 0,221$

$f_{0,025;8;7} = 4,90$

MAW: Aanvaar H_0 indien $0,221 < f_s < 4,90$
Verwerp H_0 indien $f_s \leq 0,221$ of $f_s \geq 4,90$

$f_s = \frac{0,0076}{0,0134} = 0,567$

DUS: H_0 word aanvaar.
Die variasies is wel dieselfde.

OPGAWE 10

1. Toets of die volgende steekproef uit 'n normaalverdeelde populasie kom:

   29,1 30,1 29,3 29,4 26,7 30,0 30,1 29,4.

2. 'n Vervaardiger van gloeilampe beweer dat meer as 95% van die gloeilampe wat hulle vervaardig langer as 50 uur brand. In 'n ondersoek van 100 van hierdie gloeilampe het 83 langer as 50 uur gebrand. Toets by $\alpha = 0,05$ of die vervaardiger se eis regverdig is.

3. Daar word beweer dat mieliekultivar A 'n beter opbrengs as B het. In 'n proef is die volgende opbrengste, in gram, gemeet:

Kultivar A	67,7	73,2	68,6	67,6	72,7	70,1
	69,6	68,7	70,2	69,3	73,1	
Kultivar B	66,7	66,6	69,0	67,8	67,2	66,3
	67,3	67,9	66,4			

   Aanvaar dat die steekproewe uit normaalverdeelde populasies, met gelyke variansies kom en toets of die opbrengs van A beter is as die van B by $\alpha = 0,05$.

4. Toets of die aanname van gelyke variansies in opgawe 10.3 geregverdig was by $\alpha = 0,05$.

5. Toets of die aannames van normaalverdeeldepopulasie in opgawe 10.3 geregverdig was.

6. Beskou die volgende data:

		n	$\bar{x}$	s^2
Steekproef	1	50	12,3	2,3
	2	70	15,1	3,6

   Aanvaar dat die steekproewe uit normaalverdeelde populasies kom en toets by 'n 1% betekenispeil of die gemiddeldes dieselfde is.

7. Die breeksterkte van 'n sekere soort staalkabel moet 100 ton wees. In 'n ondersoek van 9 van hierdie kabels was die gemiddelde breeksterkte 95 ton met 'n standaardafwyking van 2,5 ton. Is die kabels nog aanvaarbaar by $\alpha = 0,05$?

8. Uit 200 boute wat deur masjien A vervaardig is was 21 defektief en uit 200 boute wat masjien B vervaardig het was daar 15 defektief. Toets by $\alpha = 0,10$ of masjien B in 'n beter toestand is as A.

9. 'n Verbruiker sal 'n sekere artikel aankoop indien nie meer as 10% defektief is nie. In 'n sekere besending van 150 van die artikels was 131 aanvaarbaar. Kan die verbruiker hierdie besending aanvaar by $\alpha = 0,05$?

10. 'n Vervaardiger van sigarette beweer dat sy sigarette nie meer as 20 mg nikotien bevat nie. In 'n ondersoek van 20 sigarette is 'n gemiddelde nikotieninhoud van 20,7 mg met 'n standaardafwyking van 2,1 mg gevind. Kan die vervaardiger se eis ondersteun word by $\alpha = 0,01$?

11. Die variansie, van die lesings wat 'n sekere meetinstrument lewer, behoort nie meer as 2 mm te wees nie. In 20 lesings met die instrument is 'n variansie van 2,3 mm gevind. Toets by α = 0,05 of die instrument nog aanvaarbare resultate lewer.

12. Om te toets of battery A 'n langer leeftyd as B het is 35 batterye van elke soort geneem. A het 'n gemiddelde leeftyd van 36,3 uur met 'n standaardafwyking van 2,3 uur gehad terwyl B 'n gemiddelde leeftyd van 31,2 uur met 'n standaardafwyking van 3,4 uur gehad het. Aanvaar dat die twee steekproewe uit normaalverdeelde populasies kom en toets die bewering by α = 0,01.

13. 'n Vervaardigingsproses lewer plastiekpype wat 'n druk van 6 bar kan weerstaan. 'n Nuwe vervaardigingsproses, wat die drukweerstand van die pype moet verhoog word ondersoek. 10 pype uit die nuwe proses is ondersoek en het 'n gemiddelde drukweerstand van 6,7 bar met 'n standaardafwyking van 1,2 bar gehad. Lewer die nuwe proses sterker pype by α = 10%?

14. Beskou die volgende twee steekproewe:

A:	4,67	4,16	4,82	4,39	7,00	6,60	7,54
B:	4,34	3,81	5,21	4,62	5,98	5,79	

a. Toets of die twee steekproewe uit normaalverdeelde populasies kom. α = 0,05.

b. Toets of die variansies van die populasies dieselfde is. α = 0,05.

c. Kan die gemiddeldes van die populasies vergelyk word. Indien wel toets of die populasiegemiddeldes dieselfde is. α = 0,05.

HOOFSTUK 11

REGRESSIE ANALIESE

In Hoofstuk 6 is die berekening van die lineêre regressielyn en korrelasiekoëffisiënt bespreek. In hierdie hoofstuk sal gekyk word hoe goed pas die regressielyn die data en of die korrelasiekoëffisiënt betekenisvol is.

11.1 Ondersoek na residue

Laat y_i die waargenome y-waardes wees en $\hat{y}_i = a+bx_i$ die beraamde y-waardes.

Die verskil $e_i = y_i - \hat{y}_i$ word die residue genoem en $e_{is} = \frac{e_i}{s}$, met

$s = \sqrt{\frac{\Sigma e_i^2}{n-2}}$, die gestandaardiseerde residue.

Dit is bekend dat $e_{is} \sim n(0;1)$, so nie kan geen waarde aan die regressielyn geheg word nie.

As n klein is kan e_{is} op waarskynlikheidsgrafiekpapier uitgestip word om te toets of dit wel normaal verdeel is.

Verder kan e_{is} uitgestip word teen x en teen y, op gewone grafiekpapier. Die punte op beide hierdie grafieke behoort goed versprei om nul te wees en tussen die grense -2 en 2.

As daar punte buite hierdie grense val is dit gewoonlik beter om hulle uit die hele ontleding te laat. As daar verder enige tendense in hierdie grafiek is beteken dit dat die residue nie normaal verdeel is nie.

Voorbeeld 1

Vir die data van voorbeeld 1 hoofstuk 6 was die regressielyn $\hat{y}$ = 150,08 + 6,72x.

Die volgende tabel word nou saamgestel:

x	y	$\hat{y}$	$e=y-\hat{y}$	e^2	e_s
9	233	210,6	22,4	501,76	0,36
13	148	237,4	-89,4	7992,36	-1,44
6	181	190,4	- 9,4	88,36	-0,15
11	265	224,0	41,0	1681,00	0,66
16	293	257,6	35,4	1253,16	0,57
			0,0	11516,64	0,00

$s = \sqrt{\frac{\Sigma e_i^2}{n-2}} = \sqrt{\frac{11516,64}{3}} = 61,96$

(Let op dat $\Sigma e_i \approx 0$ en $\Sigma e_{is} \approx 0$).

Vir die normaalgrafiek van e_{is} word die volgende tabel opgestel.

i	1	2	3	4	5
e_{is}	-1,44	-0,15	0,36	0,57	0,66
(100i)/(n+1)	16,7	33,3	50,0	66,7	83,3

Sien grafiek 2. Hierdie punte lê nie naby 'n reguitlyn nie en dus kan die regressielyn nie gebruik word nie.

Voorbeeld 2

Beskou die data

x	7,0	6,3	7,2	6,0	6,6	7,0	7,4	6,5	6,2	6,7	6,5	6,8
y	15,5	15,0	18,0	13,5	15,6	16,8	17,8	16,0	13,2	14,5	13,9	15,2

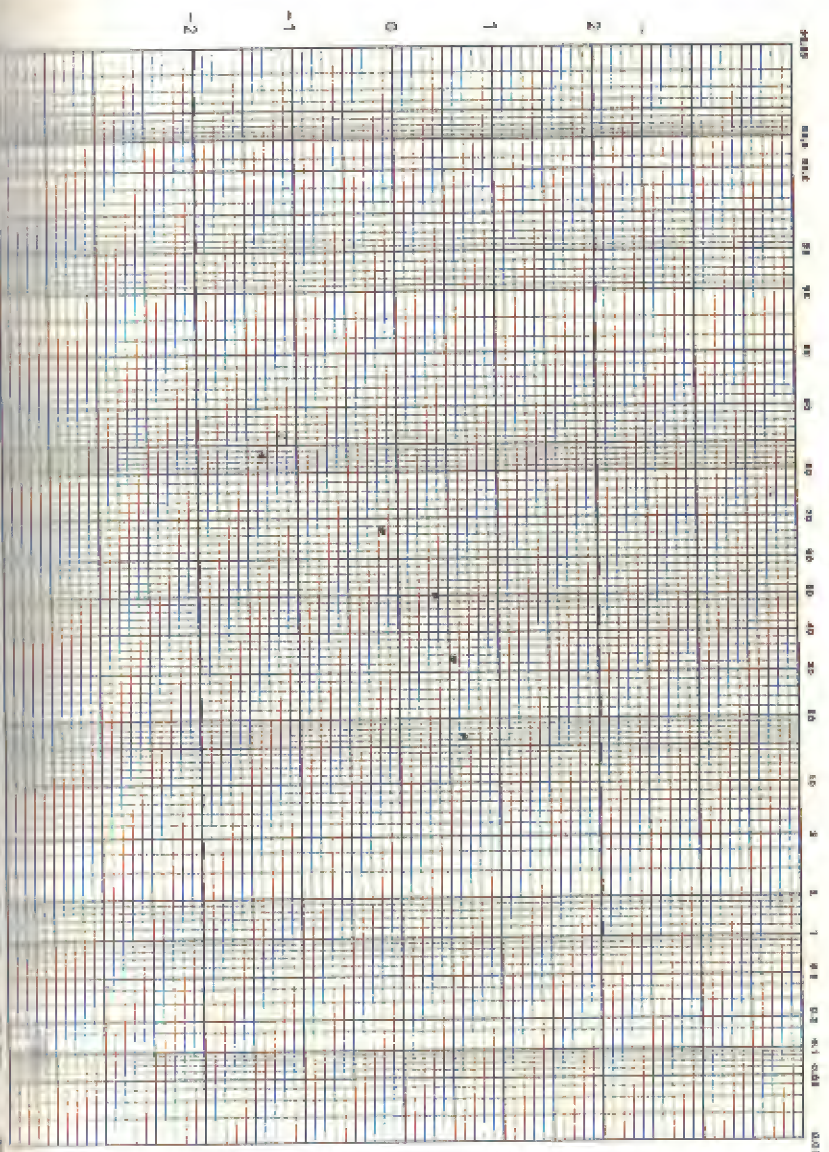
Uit die data volg dat $\hat{y} = -6{,}07 + 3{,}22x \quad r = 0{,}86$

x	y	$\hat{y}$	e	e^2	e_s
7,0	15,5	16,47	-0,97	0,94	-1,18
6,3	15,0	14,22	0,78	0,61	0,95
7,2	18,0	17,11	0,89	0,79	1,09
6,0	13,5	13,25	0,25	0,06	0,30
6,6	15,6	15,18	0,42	0,18	0,51
7,0	16,8	16,47	0,33	0,11	0,40
7,4	17,8	17,76	0,04	0,00	0,05
6,5	16,0	14,86	1,14	1,30	1,39
6,2	13,2	13,89	-0,69	0,48	-0,84
6,7	14,5	15,50	-1,00	1,00	-1,22
6,5	13,9	14,86	-0,96	0,92	-1,17
6,8	15,2	15,83	-0,63	0,39	-0,77
			-0,4	6,78	-0,49

$$s = \sqrt{\frac{\Sigma e_i^2}{n-2}} = \sqrt{\frac{6{,}78}{10}} = 0{,}82$$

Sien grafiek 3. Op beide hierdie grafieke lê die punte goed verspreid tussen die grense -2 en 2 en daar kan aanvaar word dat die lyn die data goed pas.

GRAFIEK 2



Vir die normaalgrafiek van e_{is} word die volgende tabel opgestel.

i	1	2	3	4	5
e_{is}	-1,44	-0,15	0,36	0,57	0,66
(100i)/(n+1)	16,7	33,3	50,0	66,7	83,3

Sien grafiek 2. Hierdie punte lê nie naby 'n reguitlyn nie en dus kan die regressielyn nie gebruik word nie.

Voorbeeld 2

Beskou die data

x	7,0	6,3	7,2	6,0	6,6	7,0	7,4	6,5	6,2	6,7	6,5	6,8
y	15,5	15,0	18,0	13,5	15,6	16,8	17,8	16,0	13,2	14,5	13,9	15,2

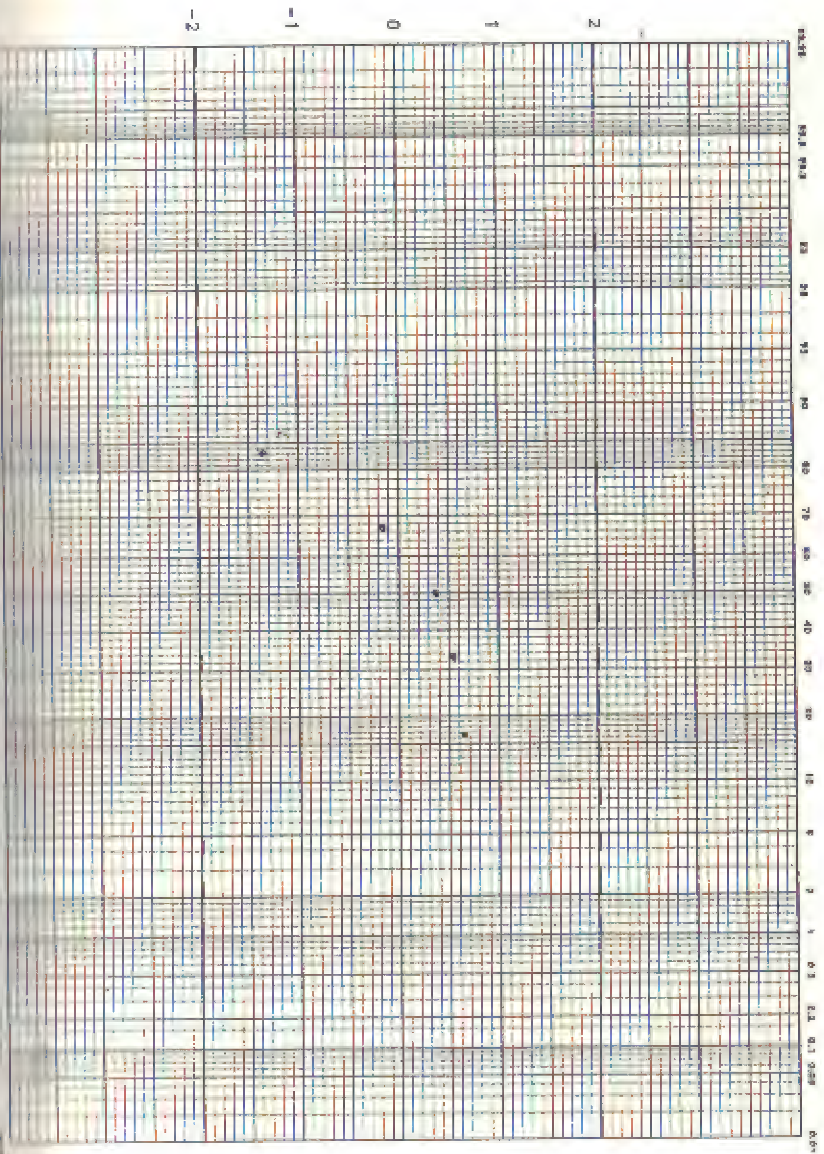
Uit die data volg dat $\hat{y} = -6,07 + 3,22x \quad r = 0,86$

x	y	$\hat{y}$	e	e^2	e_s
7,0	15,5	16,47	-0,97	0,94	-1,18
6,3	15,0	14,22	0,78	0,61	0,95
7,2	18,0	17,11	0,89	0,79	1,09
6,0	13,5	13,25	0,25	0,06	0,30
6,6	15,6	15,18	0,42	0,18	0,51
7,0	16,8	16,47	0,33	0,11	0,40
7,4	17,8	17,76	0,04	0,00	0,05
6,5	16,0	14,86	1,14	1,30	1,39
6,2	13,2	13,89	-0,69	0,48	-0,84
6,7	14,5	15,50	-1,00	1,00	-1,22
6,5	13,9	14,86	-0,96	0,92	-1,17
6,8	15,2	15,83	-0,63	0,39	-0,77
			-0,4	6,78	-0,49

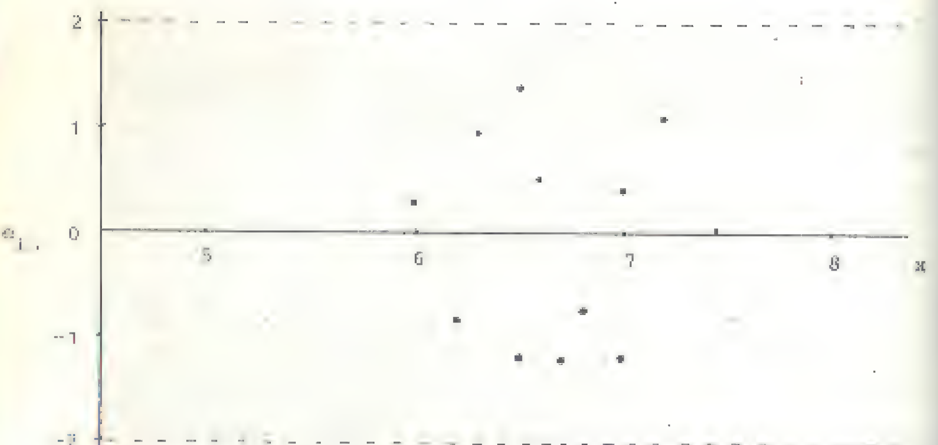
$$s = \sqrt{\frac{\Sigma e_i^2}{n-2}} = \sqrt{\frac{6,78}{10}} = 0,82$$

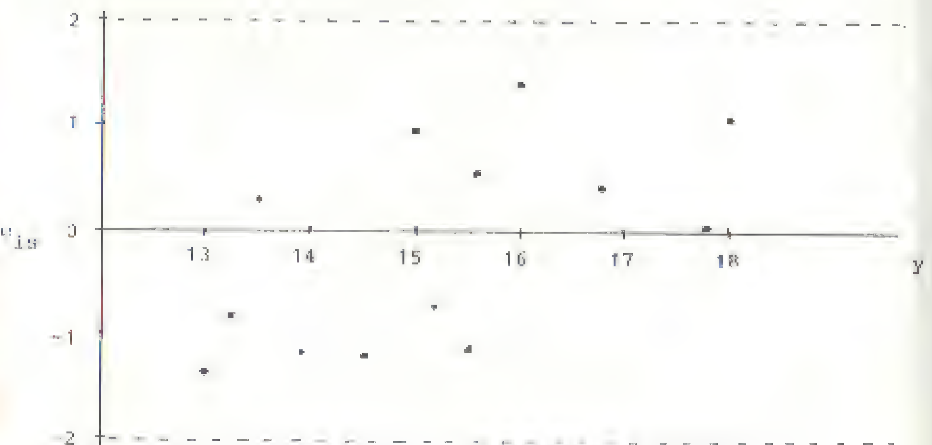
Sien grafiek 3. Op beide hierdie grafieke lê die punte goed verspreid tussen die grense -2 en 2 en daar kan aanvaar word dat die lyn die data goed pas.

GRAFIEK 2



GRAFIEK 3





Toets vir korrelasie

Nadat die aanname van normaliteit getoets is en die regressielyn was aanvaarbaar kan die korrelasiekoëffisiënt ondersoek word om te bepaal of dit betekenisvol is of nie.

Die doel van die toets is om te bepaal of die data uit 'n populasie kom waar die populasiekorrelasiekoëffisiënt, ρ(rho), nul is.

Hipoteses: $H_0: \rho = 0$

$H_1: \rho \neq 0$

Toetsstatistiek: $R_s = r$

Kritieke waarde: $R_s \sim r(n)$ tweekantig afgelees uit tabel IX, pp 14 van Stoker.

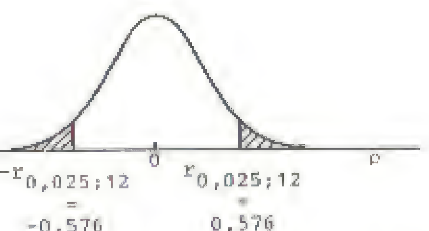
Indien H_0 aanvaar word is daar nie 'n betekenisvolle korrelasie nie en indien H_0 verwerp word is daar wel betekenisvolle korrelasie.

Voorbeeld 3

Vir voorbeeld 2 is $r = 0{,}86$ $n = 12$ en by $\alpha = 0{,}05$.

$H_0: \rho = 0$
$H_1: \rho \neq 0$

$R_s = r \sim r(n)$ onder H_0.



MAW: Aanvaar H_0 indien $-0{,}576 < r < 0{,}570$
Verwerp H_0 indien $r \leq -0{,}576$ of $r \geq 0{,}576$

$r = 0{,}86$.

DUS: H_0 word verwerp.
Die korrelasie is betekenisvol.

OPGAWE 11

1. Ondersoek die residue en toets of die korrelasiekoëffisiënt betekenisvol is by opgawes 6.1, 6.3 en 6.4. $\alpha = 0{,}05$.

2. Beskou die volgende data:

x	7,52	8,84	7,24	10,88	8,19	6,55	6,97	7,19	9,00	6,78	6,71	7,43
y	1,12	1,00	0,95	1,16	1,10	1,08	0,89	0,74	1,10	1,50	0,75	1,61

a. Pas 'n reguitlyn aan by die data en bereken die korrelasiekoëffisiënt.

b. Ondersoek die residue en toets of die korrelasiekoëffisiënt betekenisvol is.

c. Intrepreteer die resultate. $\alpha = 0{,}01$

HOOFSTUK 12

DIE ANALIESE VAN VARIANSIE

Waar daar meer as twee steekproewe is wat met mekaar vergelyk moet word gebruik ons die Analiese van Variansie prosedure om te bepaal of daar verskille tussen die steekproewe is. Indien daar wel verskille is gebruik ons meervoudige vergelykingsmetodes om te bepaal presies waar die verskille is.

12.1 EENRIGTING ANALIESE VAN VARIANSIE

Hier het ons k behandelings elk met n_i, i = 1,k waarnemings.

Laat $x_{i.}$ die behandelings totale

$\bar{x}_{i.}$ die behandelings gemiddeldes

$x_{..}$ die groottotaal en

$\bar{x}_{..}$ die algehele gemiddelde wees.

STEEKPROEF

1	2	---	k	
x_{11}	x_{21}		x_{k1}	
x_{12}	x_{22}		x_{k2}	
x_{1n_1}	x_{2n_2}		x_{kn_k}	
$x_{1.}$	$x_{2.}$	---	$x_{k.}$	$x_{..}$
$\bar{x}_{1.}$	$\bar{x}_{2.}$		$\bar{x}_{k.}$	$\bar{x}_{..}$

Laat verder $N = \sum_{i}^{k} n_i$

Die idee is nou om die totale variasie in die stel data op te deel in verskillende komponente wat ons kan toeskryf aan moontlike bronne van variasie.

Die totale variasie word gegee deur

$$SK_{totaal} = \sum_i \sum_j (x_{ij} - \bar{x}_{..})^2$$

$$= \sum_i n_i(\bar{x}_{i.} - \bar{x}_{..})^2 + \sum_i \sum_j (x_{ij} - \bar{x}_{i.})^2$$

= (totale variasie tussen behandelings) + (totale variasie binne behandelings (fout variasie))

$$= SK_{behandelings} + SK_{Fout}$$

$$= \sum_i \sum_j x_{ij}^2 - \frac{x_{..}^2}{N}$$

Die totale variasie tussen behandelings word gegee deur

$$SK_{behandelings} = \sum_i n_i(\bar{x}_{i.} - \bar{x}_{..})^2$$

$$= \sum_i x_{i.}^2/n_i - \frac{x_{..}^2}{N}$$

$$(= \frac{1}{r} \sum_i x_{i.}^2 - \frac{x_{..}^2}{N} \text{ as alle } n_i = r \text{ is})$$

Die totale variasie as gevolg van fout word gegee deur

$$SK_{Fout} = SK_{totaal} - SK_{behandelings}$$

$x_{..}^2/N$ is die korreksiefaktor (KF).

Let op dat:

Gemiddelde Som van vierkante$_{Totaal}$ = $GK_{Totaal} = SK_{Totaal}/N-1$ skat die totale variansie.

$GK_{behandelings} = SK_{behandelings}/k-1$ skat die variansie binne die behandelings.

$GK_{Fout} = SK_{Fout}/N-k$ skat die fout variansie.

Ons weet verder ook dat $Gk_{behandelings}/GK_{Fout}$ 'n F-verdeling besit met k-1 en N-k vryheidsgrade (vg.)

Die groothede wat ons hierbo bereken het word nou as volg in 'n tabel opgesom:

ANOVA TABEL

VAN VARIASIE	VG	SK	GK	F
[illegible]	k-1	$SK_{behandelings} = A$	$GK_{behandelings} = A/k-1$	$F = GK_{behandelings}/GK_{Fout}$
[illegible]	N-k	$SK_{Fout} = B$	$GK_{Fout} = B/N-k$	
[illegible]	N-1	SK_{Totaal}		

Hierdie F waarde toets ons teen 'n tabel-F waarde met (k-1) en (N-k) vryheidsgrade. (eenkantig)

Ons verwerp Ho : $\mu_1 = \mu_2 = --- \mu_k$ as $F > F_{Tabel}$.

Voorbeeld 1

Die waarnemings in die volgende tabel is die massas op ses maande van drie varkrasse. Ons wil toets of daar verskille tussen die rasse is. $\alpha = 0,05$

RAS A	RAS B	RAS C	
62	56	66	
65	58	69	
67	59	70	
70	60	71	
71	60	74	
73	63	76	
	64		
$n_1 = 6$	$n_2 = 7$	$n_3 = 6$	$N = 19$
$x_{1.} = 408$	$x_{2.} = 420$	$x_{3.} = 426$	$x_{..} = 1254$
$\bar{x}_{1.} = 68$	$\bar{x}_2 = 60$	$\bar{x}_{3.} = 71$	$\bar{x}_{..} = 66$

DIE HIPOTESES: $H_0 : \mu_1 = \mu_2 = \mu_3$

$KF = x_{..}^2/_N = 1254^2/19 = 82764$

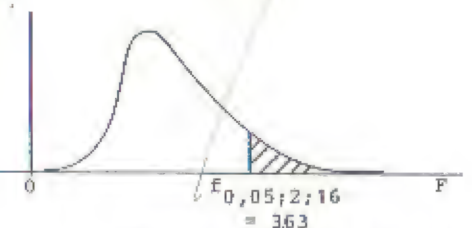
$SK_{Totaal} = \sum_i \sum_j x^2_{ij} - KF = 83384 - 82764 = 620$

$SK_{behandelings} = \sum_i x_{i.}^2/n_i - KF = 83190 - 82764 = 426$

$SK_{Fout} = SK_{Totaal} - SK_{behandelings} = 620 - 426 = 194$

ANOVA

Bron van Variasie	VG	SK	GK	f
behandelings	2	426	213	17,567
Fout	16	194	12,125	
Totaal	18	620		



MAW: Aanvaar H_0 indien $f < 3,63$.
Verwerp H_0 indien $f \geq 3,63$.

$f = 17,567$

DUS: H_0 word verwerp.
Daar is verskille tussen die rasse.

12.2 DIE TWEERIGTING ANALIESE VAN VARIANSIE

Hier het ons twee faktore A en B by k en n peile respektiewelik en r waarnemings per kombinasie. (Daar mag 'n ongelyke aantal waarnemings per kombinasie wees maar hierdie geval sal nie bespreek word nie).

FAKTOR A

FAKTOR B		Peil 1	Peil 2		Peil k	
	Peil 1	x_{111} x_{112} x_{11} ---- x_{11r}	x_{211} x_{212} x_{21} ---- x_{21r}	---- ----	x_{k11} x_{k12} $x_{k1.}$ ---- x_{k1r}	$x_{.1.}$
	Peil 2	x_{121} x_{122} $x_{12.}$ ---- x_{12r}	x_{221} x_{222} $x_{22.}$ ---- x_{22r}	---- ----	x_{k21} x_{k22} $x_{k2.}$ ---- x_{k2r}	$x_{.2.}$
	¦				¦	
	Peil n	x_{1n1} x_{1n2} $x_{1n.}$ ---- x_{1nr}	x_{2n1} x_{2n2} $x_{2n.}$ x_{2nr}	----	x_{kn1} x_{kn2} $x_{kn.}$ ---- x_{knr}	$x_{.n.}$
		$x_{1..}$	$x_{2..}$	----	$x_{k..}$	$x_{...}$

Laat $x_{.j.}$ die ry totale

$x_{i..}$ die kolom totale

$x_{ij.}$ die sel totale

$x_{...}$ die groot totaal wees.

$N = r.k.n.$

Die totale variasie in hierdie geval kan geskryf word as

variasie binne faktor A +
variasie binne faktor B +
variasie as gevolg van die interaksie tussen A en B +
variasie as gevolg van fout.

of

$$SK_{Totaal} = SK_A + SK_B + SK_{AB} + SK_{Fout}$$

Ons bereken die somme van vierkante as volg -

a. $KF = x_{...}^2/N$

b. $SK_{Totaal} = \sum_i \sum_j \sum_k x_{ijk}^2 - KF$ (N-1 vg.)

c. $SK_A = \sum_i X_{i..}^2/nr - KF$ (k-1 vg.)

d. $SK_B = \sum_i X_{.j.}^2/kr - KF$ (n-1 vg.)

e. $SK_{AxB} = \sum_i \sum_j x_{ij.}^2/r - KF$

$SK_{AB} = SK_{AxB} - SK_A - SK_B$ ((k-1)(n-1)vg.))

f. $SK_{Fout} = SK_{Totaal} - SK_A - SK_B - SK_{AB}$ (N - nk vg.)

DIE ANOVA tabel lyk as volg:

ANOVA

BRON VAN VARIASIE	VG	SK	GK	f
A	k-1	SK_A	GK_A	$f_A = GK_A/GK_{Fout}$
B	n-1	SK_B	GK_B	$f_B = GK_B/GK_{Fout}$
AB	(k-1)(n-1)	SK_{AB}	GK_{AB}	$f_{AB} = GK_{AB}/GK_{Fout}$
Fout	N-nk	SK_{Fout}	GK_{Fout}	
TOTAAL	N-1	SK_{Totaal}		

F_A, met (k-1) en (N-nk) vg. toets die hipotese

Ho: $\mu_{1a} = \mu_{2a} = ----- \mu_{ka}$

F_B, met (n-1) en (N-nk) vg. toets die hipotese

Ho: $\mu_{1b} = \mu_{2b} = ----- \mu_{nb}$

F_{AB}, met (k-1)(n-1) en (N-nk) vg. toets die hipotese

Ho: $\mu_{11} = \mu_{12} = ----- = \mu_{mk}$

Verwerp, in al die gevalle, Ho as $F > F_{tabel}$

NB. Waar r = 1 (slegs 1 waarneming per sel) kan ons nie SK_{AB} bereken nie. Hier lyk die ANOVA tabel as volg:

ANOVA

Bron van variasie	VG.	SK_A	GK	f
A	k-1	SK_A	GK_A	$f_A = GK_A/GK_{Fout}$
B	n-1	SK_B	GK_B	$f_B = GK_B/GK_{Fout}$
Fout	N-k-n+1	SK_{Fout}	GK_{Fout}	
TOTAAL	N-1	SK_{Totaal}		

Voorbeeld 2

Die effek van toediening van 2 soorte kunsmis op mielies moet ondersoek word.

Vier behandelings is dus ter sprake:

i. geen kunsmis; ii. kunsmis A; kunsmis B

iv. kunsmis A en B.

Daar was 7 waarnemings per sel en die opbrengste in kg per perseel was as volg:

KUNSMIS B		KUNSMIS A a_0	a_1	
	b_0	15 24 4 $x_{11.}$ = 91 9 6 8 25	13 17 13 $x_{21.}$ = 74 3 2 7 19	$x_{.1.}$ = 165
	b_1	5 10 2 $x_{12.}$ = 39 1 7 3 11	31 37 17 $x_{22.}$ = 189 24 19 28 33	$x_{.2.}$ = 228
		$x_{1..}$ = 130	$x_{2..}$ = 263	$x_{...}$ = 393

$k = 2 \quad n = 2 \quad r = 7 \quad N = 28$

$KF = x_{...}^2/N = \frac{393^2}{28} = 5516,04$

$SK_{Totaal} = \sum_i \sum_j \sum_k x_{ijk}^2 - KF = 8411 - 5516,04$
$= 2894,96$

$SK_A = \Sigma x_{i..}^2/nr - KF = 86069/14 - 5516,04$
$= 631,75$

$SK_B = \sum_j x_{.j.}^2/kr - KF = 79209/14 - 5516,04$
$= 141,75$

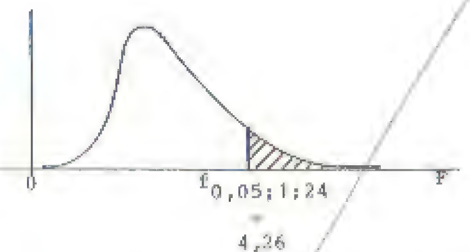
$SK_{AxB} = \sum_i \sum_j x_{ij.}^2/r - KF = 50999/7 - 5516,04$
$= 1769,53$

$SK_{AB} = SK_{AxB} - SK_A - SK_B = 1769,53 - 631,75 - 141,75$
$= 996,03$

$SK_{Fout} = SK_{Totaal} - SK_A - SK_B - SK_{AB}$
$= 2894,96 - 631,75 - 141,75 - 996,03$
$= 1125,43$

ANOVA

Bron van Variasie	VG.	SK	GK	f
A	1	631,75	631,75	13,47*
B	1	141,75	141,75	3,02
AB	1	996,03	996,03	21,24*
Fout	24	1125,43	46,89	
Totaal	27	2894,96		



MAW: $H_0 : \mu_{1a} = \mu_{2a}$ word verwerp.
$H_0 : \mu_{1b} = \mu_{2b}$ word aanvaar.
$H_0 : (\mu)_{00} = (\mu)_{01} = (\mu)_{10} = (\mu)_{11}$ word verwerp.

Kunsmis A het 'n betekenisvolle invloed en daar is 'n betekenisvolle interaksie tussen A en B.

Hierdie betekenisvolheid word aangedui deur 'n * (op 5% b.p.).

(op 1% b.p.) * (op 0,1% b.p.) langs die berekende F-waardes te plaas. (Sien ANOVA tabel).

12.3 MEERVOUDIGE VERGELYKINGS:

Indien die effekte van die peile van 'n sekere faktor van mekaar verskil moet ons nog gaan bepaal presies watter peile verskil van mekaar.

Gestel ons wil twee gemiddeldes $\bar{x}_1$ en $\bar{x}_2$ met mekaar vergelyk. Ons bereken eers, met 'n sekere formule 'n kleinste betekenisvolle verskil (KBV). Die absolute verskil tussen $\bar{x}_1$ en $\bar{x}_2$ word dan met die KBV vergelyk en as

$|\bar{x}_1 - \bar{x}_2| \geq$ KBV verskil die gemiddeldes en as

$|\bar{x}_1 - \bar{x}_2| <$ KBV verskil die gemiddeldes nie.

12.3.1 Die t-toets

$$KBV_t = \sqrt{GK_{Fout}} \quad \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad t_{\alpha/2;\nu}$$

Waar GK_{fout} die gemiddelde vierkant fout is, n_1 en n_2 die aantal waarnemings waaruit die gemiddeldes bereken is (vir beide gemiddeldes moet dit gelyk wees).

$t_{\alpha;\nu}$ is 'n tweekantige t-tabel waarde by b.p. = α met ν= vryheidsgrade vir fout.

12.3.2 Tukey se KBV

$$KBV_T \quad \sqrt{GK_{Fout}} \quad \sqrt{(\frac{1}{n_1} + \frac{1}{n_2})/2} \; q_{\alpha;k;\nu}$$

waar $q_{\alpha;k;\nu}$ 'n tweekantige "Gestudentifiseerde Variasiewydte" tabelwaarde by b.p. = α, k die aantal gemiddeldes wat vergelyk moet word en ν= vryheidsgrade van fout is.

Voorbeeld 3

Sien voorbeeld 2.

Die interaksie was betekenisvol.

KBV_t (5%)

$$KBV_t = \sqrt{46,890} \quad \sqrt{\frac{1}{7} + \frac{1}{7}} \quad 2,064 = 7,554$$

Die absolute verskille word in die volgende tabel gegee:

	$(\alpha\beta)_{00}$	$(\alpha\beta)_{10}$	$(\alpha\beta)_{01}$	$(\alpha\beta)_{11}$
$(\alpha\beta)_{00}$	-	2,429	7,429	14*
$(\alpha\beta)_{10}$	-	-	5	16,429*
$(\alpha\beta)_{01}$	-	-	-	21,429*
$(\alpha\beta)_{11}$	-	-	-	-

DUS: Die interaksie tussen

$\alpha\beta_{00}$ en $\alpha\beta_{11}$

$\alpha\beta_{10}$ en $\alpha\beta_{11}$

$\alpha\beta_{01}$ en $\alpha\beta_{11}$

is betekenisvol.

OPGAWE 12

1. Ses werkers is elk vyf kanse gegee om 'n sekere werkstuk te voltooi en die tye (min) was as volg:

Werker					
1	2	3	4	5	6
61	39	42	21	18	53
81	71	51	49	63	75
33	38	46	51	48	42
13	10	39	38	25	21
33	25	16	14	21	28

Voer 'n eenrigting ANOVA uit en bepaal of die gemiddelde tye van die werkers verskil. Gebruik KBV_t. $\alpha = 0,05$.

2. Bogenoemde data kan verdeel word in 5 dae:

		Werker					
		1	2	3	4	5	6
Dag	M	61	39	42	21	18	53
	D	81	71	51	49	63	75
	W	33	38	46	51	48	42
	D	13	10	39	38	25	21
	V	33	25	16	14	21	28

Voer 'n twee rigting ANOVA uit en bepaal of die gemiddelde per werker en per dag verskil. Gebruik KBV_t. $\alpha = 0,05$

3. 'n Twee rigting proef is uitgevoer en die resultate was as volg:

		A			
		1	2	3	4
B	1	331 312	323 322	313 319	319 291
	2	311 286	292 294	284 266	291 258
	3	280 177	170 185	250 182	253 193

Voer 'n tweerigting ANOVA uit en bepaal of A, B en AB verskil. Gebruik KBV_T. $\alpha = 0,01$.

4. Ontleed die volgende as 'n eenrigting ANOVA. Gebruik KBV_t. $\alpha = 0,05$.

Steekproef				
1	2	3	4	5
148	198	148	208	150
160	199	160	201	160
160	207	164	200	176
167	206	175	197	
175	210			

HOOFSTUK 13

GEBEURLIKHEIDSTABELLE

13.1 Die χ^2-aanpassingstoets

Dit gebeur dikwels dat 'n stel frekwensies, $o_1; o_2; \cdots; o_k$ waargeneem word uit 'n populasie en dat die teoretiesefrekwensies, $e_1; e_2; \cdots; e_k$ bereken kan word. Die vraag is nou of hierdie twee stelle frekwensies dieselfde verdeling besit.

Teoretiesefrekwensies word bereken as (waarskynlikheid) × (steekproefgrootte).

Hipoteses: H_0: waargenome- en teoretieseverdelings is dieselfde.

H_1: verdelings is nie dieselfde nie.

Toetsgrootheid: $\chi_S^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$

Kritieke waarde: $\chi_S^2 \sim \chi^2_{\alpha;(k-1)}$ (regseenkantig afgelees uit tabel IV van Stoker) as H_0 waar is.

Voorbeeld 1

Twee dobbelstene is 144 keer gerol en die volgende is waargeneem:

Som van uitkoms	2	3	4	5	6	7	8	9	10	11	12
frekwensie (o_i)	2	10	10	15	18	27	22	14	14	10	2

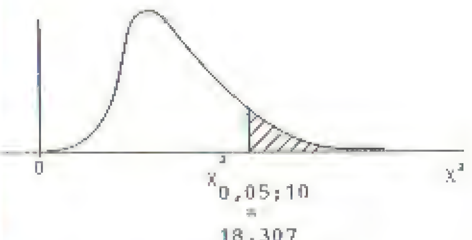
Toets by $\alpha = 0,05$ of die dobbelstene onsydig is. (dws. toets of hierdie frekwensies dieselfde is as die verwagte frekwensies.)

Som	wh. op som	e = wh × 144	o
2	1/36	4	2
3	2/36	8	10
4	3/36	12	10
5	4/36	16	15
6	5/36	20	18
7	6/36	24	27
8	5/36	20	22
9	4/36	16	14
10	3/36	12	14
11	2/36	8	10
12	1/36	4	2

H_0: verdelings is dieselfde

H_1: verdelings is nie dieselfde nie.

$\chi_S^2 = \sum \frac{(o_i - e_i)^2}{e_i} \sim \chi^2$ onder H_0.



MAW: Aanvaar H_0 indien $\chi_S^2 < 18,307$

Verwerp H_0 indien $\chi_S^2 \geq 18,307$.

$$\chi_S^2 = \frac{(2-4)^2}{4} + \ldots + \frac{(2-4)^2}{4} = 4,75$$

DUS: H_0 word aanvaar.
Die dobbelstene is onsydig.

13.2 Tweerigting-gebeurlikheidstabelle

Waargenome frekwensie kan ook in terme van tweerigtingtabelle gegee word.

		Faktor A				
		a_1	a_2	---	a_k	
Faktor B	b_1	a_{11}	a_{12}	---	a_{1k}	$a_{1\cdot}$
	b_2	a_{21}	a_{22}		a_{2k}	$a_{2\cdot}$
	⋮	⋮			⋮	⋮
	b_r	a_{r1}	a_{r2}	---	a_{rk}	$a_{r\cdot}$
		$a_{\cdot 1}$	$a_{\cdot 2}$	---	$a_{\cdot k}$	$a_{\cdot\cdot}$

Vir so 'n tweerigting-tabel word die teoretiesefrekwensies as volg bereken:

Faktor A

Faktor B	a_1	a_2	---	a_k	
b_1	$e_{11} = \frac{O_{1.} \times O_{.1}}{O_{..}}$	$e_{12} = \frac{O_{1.} \times O_{.2}}{O_{..}}$	---	$e_{1k} = \frac{O_{1.} \times O_{.k}}{O_{..}}$	$e_{1.} = O_{1.}$
⋮	⋮			⋮	⋮
b_r	$e_{r1} = \frac{O_{r.} \times O_{.1}}{O_{..}}$	$e_{r2} = \frac{O_{r.} \times O_{.2}}{O_{..}}$	---	$e_{rk} = \frac{O_{r.} \times O_{.k}}{O_{..}}$	$e_{r.} = O_{r.}$
	$e_{.1} = O_{.1}$	$e_{.2} = O_{.2}$	---	$e_{.k} = O_{.k}$	$e_{..} = O_{..}$

Toetsstatistiek: $X_S^2 = \sum_{ij} \frac{(O_{ij} - e_{ij})^2}{e_{ij}}$

Kritieke waarde: $X_S^2 \sim \chi^2_{\alpha;(k-1)(r-1)}$ (regseenkantig afgelees uit tabel IV van stoker).

as H_0 waar is.

Hipoteses: Tweerigting-gebeurlikheidstabelle kan in twee groepe verdeel word. Die toetsstatistiek en kritieke waarde by die twee groepe is dieselfde maar verskillende hipoteses word getoets.

i. Vaste groottotaal

Hier word 'n steekproef van n elemente getrek en twee veranderlikes op elke element gemeet.

Die hipotese wat getoets word is

H_0 : Die twee veranderlikes is onafhanklik

H_1 : Die twee veranderlikes is afhanklik.

ii. Vaste rytotaal (of kolomtotaal)

Hier word uit r populasies, rye, steekproewe getrek en elke element volgens k kategorieë, kolomme, geklassifiseer.

Die hipoteses wat getoets word is

H_0 : kategorieë (kolomme) is ewemoontlik by al die steekproewe

H_1 : kategorieë is nie ewemoontlik by al die steekproewe nie.

Voorbeeld 2

'n Steekproef van 100 persone is geneem en die volgende tweerigtingtabel is saamgestel

O_{ij}

		Ouderdom (jare)			
		18-25	26-35	36+	
Inkomste (B)	0-10000	12	8	5	25
	10001-20000	5	26	4	35
	20001-30000	2	7	16	25
	30001+	1	5	9	15
		20	46	34	100

Toets by $\alpha = 0,05$ of inkomste en ouderdom onafhanklik is.

H_0 : Inkomste en ouderdom is onafhanklik

H_1 : Inkomste en ouderdom is afhanklik.

e_{ij}

		Ouderdom (jare)			
		18-25	26-35	36+	
Inkomste (B)	0-10000	5,0	11,5	8,5	25
	10001-20000	7,0	16,1	11,9	35
	20001-30000	5,0	11,5	8,5	25
	30000+	3,0	6,9	5,1	15
		20	46	34	100

$$X_S^2 = \sum_{ij} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2 \text{ onder } H_0.$$

0 $\qquad \chi^2_{0,05;6} = 12,592 \qquad \chi^2$

MAW: Aanvaar H_0 indien $X_S^2 < 12,592$.

Verwerp H_0 indien $X_S^2 \geq 12,592$.

$$X_S^2 = \frac{(12 - 5,0)^2}{5,0} + \ldots + \frac{(9 - 5,1)^2}{5,1} = 39,23$$

DUS: H_0 word verwerp.
Inkomste en ouderdom is afhanklik.

Voorbeeld 3 *Hierdie is 2 sp.*

Honderd rokers en honderd nie-rokers is gevra of hulle die afgelope jaar griep gehad het of nie en die resultate was as volg:

	nie-rokers	rokers	
het griep gehad	55	74	129
het nie griep gehad	45	26	71
	100	100	200

Toets of die waarskynlikheid dieselfde is om griep te kry by rokers en nie-rokers. $\alpha = 0,01$.

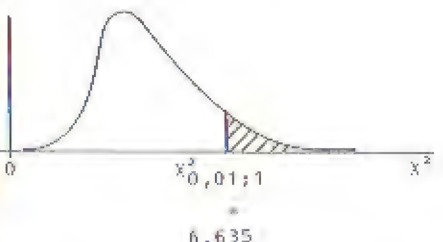
H_0 : waarskynlikheid is dieselfde

H_1 : waarskynlikheid is nie dieselfde.

e_{ij}

	nie-rokers	rokers	
het griep gehad	64,5	64,5	129
het nie griep gehad	35,5	35,5	71
	100	100	200

$$X_s^2 = \sum_{ij} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2 \text{ onder } H_0.$$



MAW: Aanvaar H_0 indien $X_s^2 < 6,635$

Verwerp H_0 indien $X_s^2 \geq 6,635$

$$X_s^2 = \frac{(55 - 64,5)^2}{64,5} + \cdots + \frac{(26 - 35,5)^2}{35,5} = 7,88$$

DUS: H_0 word verwerp.
Die waarskynlikheid om griep te kry by rokers en nie-rokers is nie dieselfde nie.

OPGAWE 13

1. Dit is bekend dat die tipes PQ, pQ, Pq en pq in die verhouding 9:3:3:1 voorkom by genetiese teelwerk. In 'n sekere eksperiment is die volgende frekwensie waargeneem

Tipe	PQ	pQ	Pq	pq
Frekwensie	170	78	71	41

Toets by $\alpha = 0,05$ of hierdie frekwensies dieselfde is as die verwagte frekwensie.

2. Dit word verwag dat in 'n sekere eksperiment A,B,C en D in dieselfde verhouding voor sal kom. Die volgende frekwensies is waargeneem.

Tipe :	A	B	C	D
Frekwensie :	43	57	49	51

Toets by $\alpha = 0,10$ of aan die verwagting voldoen is.

3. Beskou die volgende twee-rigting tabel:

		Kolomme 1	2	3	4	
rye	1	54	63	56	62	235
	2	28	30	36	26	120
	3	8	12	8	17	45
		90	105	100	105	400

Toets by $\alpha = 0,05$ of die rye en kolomme onafhanklik is.

4. 50 vroulike en 75 manlike werkers is in 'n sekere fabriek as volg geklassifiseer:

		Tipe werk: Arbeider	Klerklik	Bestuur	
Werker	Vroulik	14	27	9	50
	Manlik	20	38	17	75
		34	65	26	125

Toets by $\alpha = 0,01$ of die waarskynlikheid op die tipe werker dieselfde is by vroulike en manlike werkers.

HOOFSTUK 14

VERDELINGSVRYETOETSE

Al die hipotesetoetse wat tot nou toe bespreek is maak sekere aannames oor die populasies waaruit die steekproewe geneem is. (Bv. die populasie besit 'n normaalverdeling.) Hierdie toetse word parametriesetoetse genoem. Voordat enige van hierdie toetse gebruik kan word behoort die geldigheid van hierdie aannames eers ondersoek word.

Verdelingsvryetoetse inteendeel maak nie hierdie aannames nie en as daar getwyfel word oor die geldigheid van sekere aannames behoort 'n verdelingsvryetoets gedoen te word.

Oor die algemeen vergelyk die doeltreffendheid van verdelingsvryetoetse baie goed met die parametriesetoetse.

Slegs twee van hierdie toetse word bespreek maar daar bestaan 'n verdelingsvryetoets vir byna elke parametriesetoets.

14.1 Die Tekentoets

In §10.6.1 is 'n parametriesetoets bespreek om die hipotese $H_0: \mu_1=\mu_2$ teen die alternatief $H_1: \mu_1 \neq \mu_2$ te toets by twee klein, onafhanklike steekproewe uit normaalverdelings.

Die Tekentoets wat toets of twee steekproewe uit identies verdeelde populasies kom kan in die plek van bogenoemde toets gedoen word, met die voordeel dat die aannames nie gemaak word nie.

Hipotesis: H_0 : verdelings is dieselfde

H_1 : verdelings is nie dieselfde nie.

Toetsgrootheid: Laat x_i en y_i, $i = 1;\cdots;n$ pare waarnemings uit twee populasies wees. Laat verder $d_i = x_i - y_i$ en S_S die aantal positiewe d's.

Kritieke waarde: Die onderste kritieke waarde vir S_S word afgelees uit table XIII pp 20 van Stoker. Die boonste kritieke waarde is n-onderste kritieke waarde.

VOORBEELD 1: Die dividende (%) wat twee aandele, A en B, die afgelope 10 jaar betaal het was as volg:

Jaar	83	82	81	80	79	78	77	76	75	74
A	7,9	8,5	9,4	7,8	7,2	5,6	5,8	5,9	3,9	5,8
B	8,4	8,1	8,0	7,6	6,1	5,4	6,0	6,1	5,9	6,1

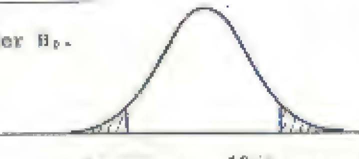
Toets by $\alpha = 0,05$ of die dividende dieselfde was.

H_0 : verdeling is dieselfde

H_1 : verdeling is nie dieselfde nie.

A = x	B = y	d = x-y	Teken van d
7,9	8,4	-0,5	-
8,5	8,1	0,4	+
9,4	8,0	1,4	+
7,8	7,6	0,2	+
7,2	6,1	1,1	+
5,6	5,4	0,2	+
5,8	6,0	-0,2	-
5,9	6,1	-0,2	-
3,9	5,9	-2,0	-
5,8	6,1	-0,3	-

S_S ~ tabel XIII pp 20 van Stoker onder H_0.



$s_{0,025;10} = 1$ $\qquad 10-s_{0,025;10} = 9$

[illegible]: Aanvaar H_0 indien $1 < s_S < 9$

Verwerp H_0 indien $s_S \leq 1$ of $S_S \geq 9$

$S_S = 5$

[illegible]: H_0 word aanvaar.

Die verdelings is dieselfde, m.a.w. die opbrengste was dieselfde dieselfde.

14.2 Die Wilcoxon rang-som toets

Hierdie is 'n verdelingsvrye toets soortgelyk aan die toets in §10.6.1 waar $H_0 : \mu_1 = \mu_2$ teen $H_1: \mu_1 < \mu_2$ getoets word.

Uit N eksperimentele eenhede word n_1 eenhede aan 'n sekere behandeling onderwerp en die orige $m_1 = N-n_1$ eenhede aan 'n kontrole. Aan die einde van die eksperiment word daar verwag dat die eenhede wat die behandeling ontvang het groter opbrengste sal lewer as die kontrole eenhede.

Gestel die opbrengste van die eksperiment is $X_1;X_2;\cdots X_{n_1}$ vir die behandeling en $Y_1;Y_2;\cdots;Y_{m_1}$ vir die kontrole. Die range $1;2;\cdots;N$ word toegeken aan $X(i)$ en $Y(i)$ sê $R_1;R_2;\cdots,R_{n_1}$ by die behandeling en $S_1;S_2;\cdots S_{m_1}$ by die kontrole. Nou word ΣR_i en ΣS_i bereken. (Let op dat $\Sigma R_i + \Sigma S_i = \frac{1}{2}N(N+1)$).

Hipotesis: H_0 : geen behandelingseffek

H_1 : wel behandelingseffek

Toetstatistiek: $W_S = \Sigma S_i - \frac{1}{2}m_1(m_1+1)$.

Kritieke waarde: Die toets is altyd linkseenkantig en die kritieke waarde word direk afgelees uit tabel XV pp 22 van Stoker by betekenispeil α en $n = \text{mak}(n_1;m_1)$ en $m = \min(n_1;m_1)$.

Voorbeeld 2: Dit word verwag dat 'n nuwe wiskunde onderrigtegniek die prestasie van studente sal verhoog. Om dit te toets is 'n groep van 20 studente gekies en uit hulle het 10 die nuwe onderrig metode ontvang en die orige 10 'n standaard metode. Een van die studente by die standaardmetode het gedurende die kursus gestaak. Die finale uitslae (%) was as volg.

Nuut	60	72	68	81	89	92	45	58	67	74
Standaard	48	87	91	56	64	51	54	66	65	

Toets by $\alpha = 0{,}01$ of die nuwe metode wel beter gevaar het.

Hipotesis: H_0 : geen behandelingseffek

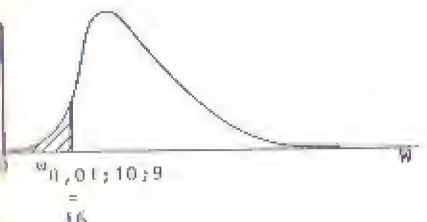
H_1 : wel behandelingseffek

X_i=nuut	Y_i=standaard	$X_{(i)}$	$Y_{(i)}$	R_i	S_i
60	48	45	48	1	2
72	87	58	51	6	3
68	91	60	54	7	4
81	56	67	56	11	5
89	64	68	64	12	8
92	51	72	65	13	9
45	54	74	66	14	10
58	66	81	87	15	16
67	65	89	91	17	18
74		92		19	
$n_1=10$	$m_1=9$			115	75

$\Sigma R_i + \Sigma S_i = 190 \qquad \frac{1}{2}N(N+1) = \frac{1}{2}(19)(20) = 190$

$W_S = \Sigma S_i - \frac{1}{2}m(m+1) \sim W$ onder H_0.

$n = 10 \quad m = 9$



DAN: Aanvaar H_0 indien $w_s > 16$.

Verwerp H_0 indien $w_s \leqslant 16$.

$w_s = 75 - \frac{1}{2}(9)(10) = 30$

DUS: H_0 word aanvaar.
Die metode lewer nie beter resultate nie.

Waar daar gelyke waarnemings is word gemiddelde range toegeken.

Voorbeeld 3

X_i	Y_i	$X_{(i)}$	$Y_{(i)}$	R_i	S_i
[illegible]	8	5	6	1	2,5
[illegible]	8	6	7	2,5	4
[illegible]	7	8	8	6	6
[illegible]	12	9	8	8	6
	6		12		9
				17,5	27,5

$\frac{1}{2}N(N+1) = \frac{1}{2}(9)(10) = 45$

$17{,}5 + 27{,}5 = 45$.

15.2 Ongeweegde saamgestelde prysindeks

Vir 'n groep artikels is die ongeweegde saamgestelde prysindeks

$$P_S = \frac{\Sigma p_n}{\Sigma p_o} \times 100$$

Vir artikels A, B en C in die voorbeeld is

$$P_S = \frac{12,73}{12,32} \times 100 = 103.$$

(In totaal het die pryse van die drie artikels met 3% gestyg.)

15.3 Geweegde prysindekse

Dit is vanselfsprekend dat alle artikels nie ewe veel nut vir die verbruiker het nie. Mieliemeel het byvoorbeeld 'n baie groter nut as vleissmeer.

Die nut wat 'n artikel het kan gekwantifiseer word en dan word dit die gewig van die artikel genoem.

Voorbeeld 2

Artikel	Gewig	1983 Hoeveelheid (h_o)	1983 Prys (p_o)	1984 Hoeveelheid (h_n)	1984 Prys (p_n)
A	1	100	1,35	121	2,12
B	5	63	0,72	50	0,79
C	2	9	10,25	12	9,82
	8	172	12,32	183	12,73

15.3.1 Geweegde saamgestelde prysindeks

$$P_{gs} = \frac{\Sigma p_n g}{\Sigma p_o g} \times 100$$

Vir die voorbeeld:

$$P_{gs} = \frac{(2,12)(1) + (0,79)(5) + (9,82)(2)}{(1,35)(1) + (0,72)(5) + (10,25)(2)} \times 100 = \frac{25,71}{25,45} \times 100$$

$= 101$

15.4 Laspeyres se prysindeks

$$P_L = \frac{\Sigma p_n h_o}{\Sigma p_o h_o} \times 100$$

Vir die voorbeeld is

$$P_L = \frac{(2,12)(100) + (0,79)(63) + (9,82)(9)}{(1,35)(100) + (0,72)(63) + (10,25)(9)} \times 100 = \frac{350,15}{272,61} \times 100$$

$= 128$

15.5 Paasche se prysindeks

$$P_P = \frac{\Sigma p_n h_n}{\Sigma p_o h_n} \times 100$$

Vir die voorbeeld is

$$P_P = \frac{(2,12)(121) + (0,79)(50) + (9,82)(12)}{(1,35)(121) + (0,72)(50) + (10,25)(12)} \times 100 = \frac{413,86}{332,35} \times 100$$

$= 128$

15.6 Fischer se prysindeks

$$P_F = \sqrt{P_L \cdot P_P}$$

Vir die voorbeeld:

$$P_F = \sqrt{(128)(128)} = 128$$

15.7 Drobisch se prysindeks

$$P_D = \frac{(P_L + P_P)}{2}$$

Vir die voorbeeld:

$$P_D = \frac{128 + 128}{2} = 128$$

15.8 Verandering van die basisjaar van prysindekse

As 'n stel indekse met 'n sekere basisjaar bereken is, is dit soms nodig om hierdie basisjaar te verander:

$$\text{Nuwe indeks} = \frac{\text{ou indeks}}{\text{basisjaar indeks}} \times 100.$$

Voorbeeld 3

Die volgende tabel gee die prysindekse van 'n groep artikels vir die afgelope 7 jaar, met 1976 as basisjaar.

Jaar	Indeks
1977	82
1978	93
1979	110
1980	111
1981	109
1982	121
1983	113

As 1978 die basisjaar gemaak moet word is die nuwe indekse:

Jaar	Indeks
1977	$\frac{82}{93} \times 100 = 88$
1978	$\frac{93}{93} \times 100 = 100$
1979	$\frac{110}{93} \times 100 = 118$
1980	$\frac{111}{93} \times 100 = 119$
1981	$\frac{109}{93} \times 100 = 117$
1982	$\frac{121}{93} \times 100 = 130$
1983	$\frac{113}{93} \times 100 = 122$

(Let op dat die indeks vir die nuwe basisjaar 100 is.)

OPGAWE 15

1. Beskou die volgende tabel:

	Gewig	1981 Prys	1981 Hoeveelheid	1982 Prys	1982 Hoeveelheid	1983 Prys	1983 Hoeveelheid
Botter	2	1,73	64	1,95	75	1,25	120
Melk	8	0,51	127	0,62	136	0,84	172
Kaas	4	1,23	73	1,75	72	2,01	74

a. Met 1981 as basisjaar bereken enkelvoudige ongeweegde prysindekse vir botter, melk en kaas in 1982 en 1983.

b. Met 1982 as basisjaar bereken ongeweegde saamgestelde, Laspeyres en Paasche se prysindeks vir 1983.

c. Met 1981 as basisjaar bereken die geweegde saamgestelde prysindeks vir 1982.

2. Beskou die volgende tabel:

	Gewig	1951 Prys	1951 hoeveelheid	1952 Prys	1952 Hoeveelheid
Koring	4	125	93	137	97
Mielies	4	114	172	121	196
Sorgum	1	96	21	93	36

Met 1951 as basisjaar bereken Laspeyres, Paasche, Fischer en Drobisch se prysindekse. en hoeveelheidsindeks.

3. Beskou die volgende tabel van prysindeks (1915 was as basisjaar geneem).

Jaar	Indeks
1920	85
1925	93
1930	92
1935	97
1940	101
1945	99
1950	105
1955	110
1960	117
1965	121
1970	113
1975	101
1980	107

Bereken nuwe indekse deur 1920, 1930, 1940, 1950, 1960, 1970 en 1980 basisjare te maak.

HOOFSTUK 16

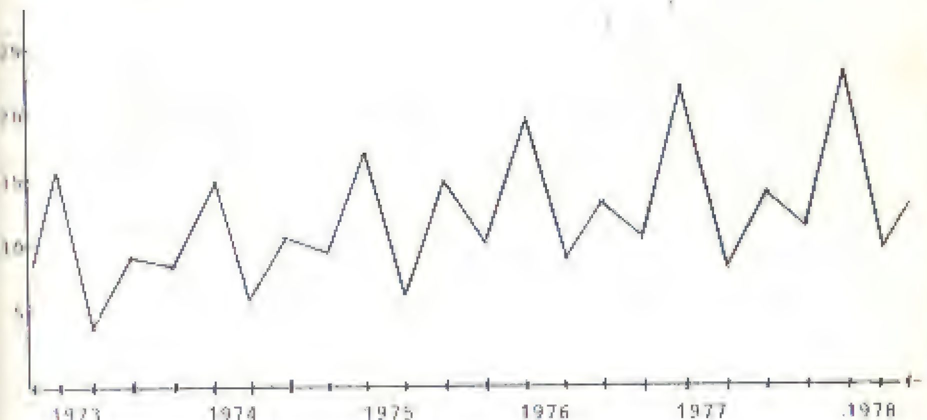
TYDREEKSE

'n Tydreeks is 'n stel data wat oor verskillende tye waargeneem is. Die doel van 'n tydreeksontleding is om die aard van die variasie te bestudeer en om voorspellings vir die toekoms te maak.

VOORBEELD

Jaar	Kwartaal	y = Verkope van ABC (R10000)
1972	3	7,4
	4	15,1
1973	1	4,7
	2	8,7
	3	7,7
	4	14,1
1974	1	6,3
	2	10,6
	3	8,6
	4	16,7
1975	1	6,7
	2	14,5
	3	10,2
	4	18,9
1976	1	8,7
	2	13,1
	3	10,8
	4	21,4
1977	1	8,2
	2	13,9
	3	11,2
	4	22,2
1978	1	9,1
	2	13,5

VERKOPE VAN ABC



Die mees algemene tydreeksmodel is die multiplikatiewe-model wat stel dat opbrengs = Langtermyneffek × Seisoenseffek × konjunktuurskommeling × toevallige variasie.

OF

$y = L \times S \times K \times V$

Die langtermyneffek (L) is die algemene tendens wat in tydreekse voorkom. Indien hierdie effek bereken word word dit in dieselfde eenhede as y uitgedruk.

Waar 'n tydreeks oor kwartale, maande of dae gaan vir 'n paar jaar is daar baie keer patrone oor die seisoene of seisoenseffekte (S) teenwoordig. As S bereken word word dit as 'n indeks uitgedruk.

Konjunktuurskommeling (K) is baie soortgelyk as seisoenseffekte maar die periodes waaroor dit plaasvind is nie konstant nie. Hierdie skommeling word toegeskryf aan faktore soos natuurrampe en politiekeklimate. As K bereken word word dit ook as 'n indeks uitgedruk.

In 'n tydreeksmodel moet daar ook voorsiening gemaak word vir toevallige variasie (V) in y. Hierdie komponent maak voorsiening vir metingsfoute en ander toevallige verskynsels. As V bereken word word dit ook as 'n indeks uitgedruk.

Enige komponent in die tydreeksmodel kan met deling verwyder word:

$$y = L\,S\,K\,V$$

$$\frac{y}{S} = \frac{L\,S\,K\,V}{S} = L\,K\,V$$

$$\frac{L\,K\,V}{L} = K\,V \text{ ens.}$$

In die grafiese voorstelling om ABC se data is dit baie duidelik dat daar 'n langtermyn en seisoenseffekte teenwoordig is.

16.1 Berekening van Langtermyneffekte

In al die gevalle wat in hierdie boek bespreek sal word kan die langtermyneffek met 'n reguitlyn benader word.

'n Metode om hierdie lyn te bereken is gewone lineêre regressie waar y as die afhanklike en tyd as onafhanklike veranderlike gekies word.

Die tye word eers geskaal sodat $\Sigma t_i = 0$. Dit kan gedoen word deur die tye ···-3; -2; -1; 0; 1; 2; 3; ··· te nommer as n onewe is en
···· -7; -5; -3; -1; 1; 3; 5; 7 ··· te nommer as n ewe is.

Nou is $L = \hat{y} = a + bt$ waar

$$b = \frac{\Sigma y_i t_i - n\bar{y}\bar{t}}{\Sigma t_i^2 - n\bar{t}^2} = \frac{\Sigma y_i t_i}{\Sigma t_i^2} \quad \text{as } \Sigma t_i = 0.$$

$$a = \bar{y} - b\bar{t} = \bar{y} \qquad \text{as } \Sigma t_i = 0.$$

Elke waarde van t_i word in die lyn vervang om $\hat{y}_i = L_i$ te bereken.

VOORBEELD (vervolg)

Jare	Kwartale	y: (R10000)	t_i	$\hat{y}_i = L$ (R10000)
1972	3	7,4	-23	8,8
	4	15,1	-21	9,0
1973	1	4,7	-19	9,3
	2	8,7	-17	9,6
	3	7,7	-15	9,8
	4	14,1	-13	10,1
1974	1	6,3	-11	10,3
	2	10,6	- 9	10,6
	3	8,6	- 7	10,9
	4	16,7	- 5	11,1
1975	1	6,7	- 3	11,4
	2	14,5	- 1	11,6
	3	10,2	1	11,9
	4	18,9	3	12,2
1976	1	8,7	5	12,4
	2	13,1	7	12,7
	3	10,8	9	12,9
	4	21,4	11	13,2
1977	1	8,2	13	13,5
	2	13,9	15	13,7
	3	11,2	17	14,0
	4	22,2	19	14,2
1978	1	9,1	21	14,5
	2	13,5	23	14,8 .

$\Sigma y_i t_i = 597,3 \quad \Sigma t_i^2 = 4600,0 \quad \Sigma t_i = 0 \quad \bar{y} = 11,76$

$a = 11,76 \quad b = 0,13 \quad \hat{y} = 11,76 + 0,13t$

16.2 BEWEGENDE GEMIDDELDES

Bewegende Gemiddeldes is 'n reeks gemiddeldes wat uit 'n data reeks bereken word deur telkens een ou term weg te laat en een nuwe term by te voeg en dan gemiddeldes te bereken.

Byvoorbeeld: vir die data 8 5 9 7 6 4 is die 3-term bewegende gemiddeldes:

$\frac{8+5+9}{3} = 7,3 \quad \frac{5+9+7}{3} = 7,0 \quad \frac{9+7+6}{3} = 7,3 \quad \frac{7+6+4}{3} = 5,7$;

en die 4-term bewegende gemiddeldes:

$\frac{8+5+9+7}{4} = 7,3 \quad \frac{5+9+7+6}{4} = 6,8 \quad \frac{9+7+6+4}{4} = 6,5$

Bewegende gemiddeldes het die nadeel dat daar minder terme is as in die oorspronklike data.

16.3 Berekening van Seisoensindekse

As bewegende gemiddeldes by 'n tydreeks bereken word het dit die effek dat die reeks glad gestryk word deurdat die seisoenseffekte en toevallig variasie verwyder word.

Beskou ons die model y = L S K V en bereken bewegende gemiddelde is die resultaat dat slegs LK oorbly en

$$\frac{y}{L\ K} = \left(\frac{y}{\text{bew.gem.}}\right) = \frac{L\ S\ K\ V}{L\ K} = SV$$

(Vir kwartaallikse data word 4-term bewegende gemiddeldes bereken en vir maandelikse data 12-term bewegende gemiddeldes.)

Waar ewe-terme bewegende gemiddeldes bereken word is dit nodig om hierdie gemiddeldes te sentreer. Dit word gedoen deur paarsgewys gemiddeldes te bereken.

Jaar	Kwartaal	Y	4-term totale	4-term bew. gem.	gesentreerde 4-term bew.gem. = L K	$\frac{Y}{L\ K}$ = S V
1972	3	7,4				
	4	15,1	35,9	8,98		
1973	1	4,7	36,2	9,05	9,02	0,52
	2	8,7	35,2	8,80	8,93	0,97
	3	7,7	36,8	9,20	9,00	0,86
	4	14,1	38,7	9,68	9,44	1,49
1974	1	6,3	39,6	9,90	9,79	0,64
	2	10,6	42,2	10,55	10,23	1,04
	3	8,6	42,6	10,65	10,60	0,81
	4	16,7	46,5	11,63	11,14	1,50
1975	1	6,7	48,1	12,03	11,83	0,57
	2	14,5	50,3	12,58	12,31	1,18
	3	10,2	52,3	13,08	12,83	0,80
	4	18,9	50,9	12,73	12,91	1,46
1976	1	8,7	51,5	12,88	12,81	0,68
	2	13,1	54,0	13,50	13,19	0,99
	3	10,8	53,5	13,38	13,44	0,80
	4	21,4	54,3	13,58	13,48	1,59
1977	1	8,2	54,7	13,68	13,63	0,60
	2	13,9	55,5	13,88	13,78	1,01
	3	11,2	56,4	14,10	13,99	0,80
	4	22,2	56,0	14,00	14,05	1,58
1978	1	9,1				
	2	13,5				

Om nou die seisoensindekse te bereken word 'n kwartaallikse- of maandeliksetabel van S V saamgestel. Uit hierdie tabel word die gemiddelde per kwartaal of maand bereken nadat die grootste en kleinste waarnemings weggelaat is. Laasgenoemde word gemoderniseerde gemiddeldes genoem.

Die onaangepaste seisoensindekse is die gemoderniseerde gemiddeldes x 100.

Om die onaangepaste seisoensindeks te laat optel na 400 vir kwartaallikse data en na 1200 vir maandelikse data word elke seisoensindeks gedeel met die totaal van al die seisoensindeks en gemaal met 400 of 1200. Dit word die aangepaste seisoensindekse genoem. Verder word die aangepaste seisoensindekse ook as 'n verhouding uitgedruk.

VOORBEELD (vervolg)

KWARTAALIKSE TABEL VAN SV

Jaar	Kwartaal 1	2	3	4	Totaal
1973	0,52	0,97	0,86	1,49	
1974	0,64	1,04	0,81	1,50	
1975	0,57	1,18	0,80	1,46	
1976	0,68	0,99	0,80	1,59	
1977	0,60	1,01	0,80	1,58	
gemoderniseerde totale	1,81	3,04	2,41	4,57	
gemoderniseerde gemiddeldes	0,603	1,013	0,803	1,523	
onaangepaste seisoens-indekse %	60,3	101,3	80,3	152,3	394,2
aangepaste seisoens-indekse %	61,2	102,0	81,5	154,5	400,0
aangepaste seisoens-indeks as verhouding	0,612	1,028	0,815	1,545	4

Nou kan die seisoenseffek uit die tydreeks verwyder word deur slegs met die betrokke seisoensindeks (as verhouding) te deel:

$$\frac{Y}{S} = \frac{L\ S\ K\ V}{S} = L\ K\ V$$

$$\frac{S\ V}{S} = V \quad \text{ens.}$$

VOORBEELD (vervolg)

Jaar	Kwartaal	Y	L	SV	$\hat{S}$	$\frac{SV}{\hat{S}} = V$	$\frac{Y}{LSV} = K$
1972	3	7,4	8,8		0,815		
	4	15,1	9,0		1,545		
1973	1	4,7	9,3	0,52	0,612	0,850	0,972
	2	8,7	9,6	0,97	1,028	0,944	0,934
	3	7,7	9,8	0,86	0,815	1,055	0,914
	4	14,1	10,1	1,49	1,545	0,964	0,937
1974	1	6,3	10,3	0,64	0,612	1,046	0,956
	2	10,6	10,6	1,04	1,028	1,012	0,962
	3	8,6	10,9	0,81	0,815	0,994	0,974
	4	16,7	11,1	1,50	1,545	0,971	1,003
1975	1	6,7	11,4	0,57	0,612	0,931	1,031
	2	14,5	11,6	1,18	1,028	1,148	1,059
	3	10,2	11,9	0,80	0,815	0,982	1,071
	4	18,9	12,2	1,46	1,545	0,945	1,061
1976	1	8,7	12,4	0,68	0,612	1,111	1,007
	2	13,1	12,7	0,99	1,028	0,963	1,042
	3	10,8	12,9	0,80	0,815	0,982	1,047
	4	21,4	13,2	1,59	1,545	1,029	1,020
1977	1	8,2	13,5	0,60	0,612	0,980	1,005
	2	13,9	13,7	1,01	1,028	0,982	1,005
	3	11,2	14,0	0,80	0,815	0,982	1,000
	4	22,2	14,2	1,58	1,545	1,023	0,989
1978	1	9,1	14,5		0,612		
	2	13,5	14,8		1,028		

16.4 Vooruitskatting

In die meeste gevalle word 'n tydreeksanaliese gedoen om vooruitskattings te maak.

Die regte waardes van t word in die langtermyn lyn vervang. Nou word hierdie vooruitskatting aangepas vir die betrokke seisoensindeks (vermenigvuldig dit met die seisoensindeks). Omdat die toevallige variasie onvoorspelbaar is word dit uit vooruitskatting geïgnoreer. Konjunktuurindekse kan ook in berekening gebring word mits die tydreeks lank is sodat 'n grafiese vooruitskatting van die konjunktuurmiddels gemaak kan word, as dit nie die geval is nie word die konjunktuur ook uit die vooruitskatting gelaat.

VOORBEELD (vervolg)

Die vooruitskattings vir die volgende vier kwartale word gemaak. (Konjunktuur word geïgnoreer.)

Jaar	Kwartaal	t	L	S	Vooruitskatting
1978	3	25	15,01	0,815	12,2
	4	27	15,27	1,545	23,6
1979	1	29	15,53	0,612	9,5
	2	31	15,79	1,028	16,2

OPGAWE 16

Ontleed die volgende tydreekse en maak vier vooruitskattings:

a.

Jaar	Maand	Waarde	Jaar	Maand	Waarde
1980	Jan	219	1982	Jan	196
	Feb	218		Feb	197
	Maart	221		Maart	201
	April	216		April	201
	Mei	213		Mei	202
	Junie	211		Junie	198
	Julie	217		Julie	195
	Aug	211		Aug	199
	Sept	217		Sept	200
	Okt	213		Okt	201
	Nov	205		Nov	204
	Des	206		Des	205
1981	Jan	206	1983	Jan	203
	Feb	207		Feb	204
	Maart	205		Maart	201
	April	203		April	202
	Mei	200		Mei	202
	Junie	203		Junie	201
	Julie	200		Julie	200
	Aug	199		Aug	204
	Sept	201		Sept	200
	Okt	200		Okt	195
	Nov	195		Nov	198
	Des	193		Des	197

b.

Jaar	Termyn	Waarde
1950	1	55,9
	2	28,1
	3	26,5
1951	1	60,1
	2	34,0
	3	33,6
1952	1	67,9
	2	40,3
	3	36,7
1953	1	77,2
	2	42,8
	3	40,4
1954	1	86,8
	2	35,9
	3	46,6
1955	1	86,7
	2	50,9
	3	40,8
1956	1	85,2
	2	46,8
	3	41,0
1957	1	80,7
	2	53,1
	3	44,6
1958	1	82,0
	2	55,5
	3	47,4
1959	1	82,3
	2	55,2
	3	49,8

c.

Jaar	Semester	Waarde
1951	1	28,2
	2	38,1
1952	1	38,3
	2	29,8
1953	1	28,5
	2	27,8
1954	1	28,7
	2	39,9
1955	1	29,3
	2	37,5
1956	1	31,7
	2	37,5
1957	1	31,1
	2	28,5
1958	1	32,0
	2	29,8
1959	1	29,0
	2	38,4
1960	1	38,5
	2	56,1
1961	1	42,3
	2	40,0
1962	1	42,7
	2	52,2
1963	1	51,5
	2	60,8
1964	1	56,5
	2	50,3
1965	1	41,7
	2	43,2
1966	1	54,6
	2	46,1

HOOFSTUK 17

FINANSIëLE BEREKENINGE

17.1 Rentebegrippe

17.1.1 Tydwaarde van geld

Dit is bekend dat R100,00 vandag baie minder werd is as vyf jaar gelede. Verder sal R100,00 wat vandag belê word teen 10% rente oor een jaar R110,00 werd wees.

Om die verskillende tipes rente te bestudeer is dit nodig om eers sekere grondbegrippe te definieer:

17.1.1.1 Kapitaal (K)

Dit is 'n gegewe bedrag kontant in een of ander geldeenheid.

17.1.1.2 Rente (r)

Dit is die vergoeding wat op 'n kapitale bedrag verdien word vir die gebruik van die kapitaal. Rente word gewoonlik jaarliks, halfjaarliks, kwartaalliks of maandeliks betaal.

17.1.1.3 Termyn (n)

Gewoonlik word die kapitale bedrag vir 'n sekere termyn gebruik en rente vir daardie termyn bereken. Die termyn verwys na die totale aantal periodes, of dele van periodes, wat die kapitaal gebruik word.

17.1.1.4 Eindbedrag (S_n)

Dit is 'n bedrag gelyk aan die oorspronklike kapitaal plus die rente wat dit verdien het aan die einde van 'n sekere periode.

17.1.1.5 Rentekoers (i)

Dit is die rente per 100 eenhede kapitaal per periode Byvoorbeeld 6½% per jaar wat dieselfde is as 0,065 per eenheid per jaar.

17.1.2 Enkelvoudige rente

Dit is rente wat slegs op die oorspronklike kapitale bedrag bereken word en aan die einde van gelyke tydsintervalle betaal word.

'n Formule vir die berekening van die slotbedrag (S_n) na n periodes as die kapitale bedrag K is en die enkelvoudige rentekoers i% per periode, is

$$S_n = K + Kn\frac{i}{100}$$

Voorbeeld 1

Die slotbedrag, as R220,00 vir 9 maande teen 'n enkelvoudige rentekoers van 8% p.j. belê word is:

$$S_n = K + Kni/100$$
$$= 220 + (220(\frac{9}{12})8)/100$$
$$= 220 + 13,2$$
$$= R233,20$$

17.1.3 Saamgestelde Rente

Dit is rente wat op die oorspronklike kapitale bedrag plus vorige rentebedrae bereken word.

Die formule vir die berekening van die slot bedrag (S_n) as 'n kapitale bedrag K belê is vir n periodes teen 'n saamgestelde rentekoers van i% per periode is:

$$S_n = K(1 + \frac{i}{100})^n$$

(Let op dat i die rentekoers per periode is, as dit nie die geval is nie moet die koers per periode bereken word. Byvoorbeeld 12% p.j. = 3% per kwartaal.)

Voorbeeld 2

R1 200,00 word vir 'n periode van 2 jaar teen 'n saamgestelde rentekoers van 14% p.j. belê. Die rente word halfjaarliks saamgestel. Bereken die eindbedrag.

14% p.j. = 7% per halfjaar.

2 jaar = 4 halfjaar periodes.

$$S_n = K(1 + \frac{i}{100})^n$$
$$= 1200(1 + \frac{7}{100})^4$$
$$= R1\ 572,96$$

17.1.4 Nominale en effektiewe rentekoerse

Nominale rentekoers is die gekwoteerde rentekoers bv: 14% rente per jaar jaarliks saamgestel of 12% rente per jaar maandeliks saamgestel.

As R250,00 byvoorbeeld teen 'n rentekoers van 10% p.j., kwartaalliks saamgestel belê word is die slot bedrag na 1 jaar:

$$S_n = 250(1 + \frac{2,5}{100})^4 = R275,95.$$

Die rente wat verdien is, is

R275,95 - R250,00 = R25,95

of $\frac{25,95}{250,00} \times 100 = 10,38$

Hierdie is die effektiewe rentekoers per jaar.

17.2 Annuïteite

'n Annuïteit word gedefinieer as 'n reeks betalings wat met gelyke tydsintervalle oor 'n sekere termyn gemaak word. Byvoorbeeld: 'n persoon belê R100,00 per maand vir vyf jaar teen 12% p.j.

17.2.1 Agternabetaalde annuïteite (gewone-)

Agternabetaalbare of gewone annuïteite is as die betalings aan die einde van die tydsintervalle gemaak word en vooruitbetaalbareannuïteite is as die betaling aan die begin van 'n tydsinterval gemaak word.

As vir 'n gewone annuïteit S_n die eindbedrag, p die betalings per termyn, i die % rente per termyn en n die aantal termyne is:

$$S_n = p\left[\frac{(1+\frac{i}{100})^n-1}{\frac{i}{100}}\right]$$

of $$p = \frac{\frac{i}{100}S_n}{(1+\frac{i}{100})^n-1}$$

of $$n = \frac{\log(\frac{iS_n}{100p}+1)}{\log(1+\frac{i}{100})}$$

Voorbeeld 3: 'n Persoon belê R200,00 per maand teen 12% rente p.j. in 'n gewone annuïteit vir 10 jaar. Bereken die eindbedrag.

$$i = \frac{12\%}{12} = 1\%$$

$$n = 10 \times 12 = 120$$

$$S_n = p\left[\frac{(1+\frac{i}{100})^n-1}{\frac{i}{100}}\right]$$

$$= 200\left[\frac{(1+\frac{1}{100})^{120}-1}{\frac{1}{100}}\right]$$

$$= 46007,74$$

Voorbeeld 4

Hoe lank moet 'n persoon maandeliks 'n bedrag van R100,00 vir 'n gewone annuïteit belê teen 14% rente p.j. sodat die eindbedrag R5 000,00 is?

$$n = \frac{\log(\frac{14.5000}{12.100.100}+1)}{\log(1+\frac{14}{12.100})}$$

$$= 39,6$$

$$\simeq 40 \text{ maande.}$$

17.2.2 Vooruitbetaalde annuïteite

Vir 'n vooruitbetaalde annuïteit is die eindbedrag $\ddot{S}_n = S_{n+1} - p$ waar S_{n+1} die eindbedrag van 'n gewone annuïteit is na n+1 periodes.

Voorbeeld 5

'n Persoon belê R2 500,00 halfjaarliks in 'n vooruitbetaalde annuïteit teen 16% p.j. vir 4 jaar. Bereken die eindbedrag.

$$n = 8$$

$$i = \frac{16\%}{2} = 8\%$$

$$\ddot{S}_n = S_{n+1} - p$$

$$\ddot{S}_8 = 2\,500\left[\frac{(1+\frac{8}{100})^9-1}{\frac{8}{100}}\right] - 2\,500$$

$$= 28718,89$$

17.3 Terugbetaling van lenings

As 'n persoon 'n kapitale bedrag K leen en dit terug betaal in n periodes teen 'n rentekoers van i% per periode en die paaiement per periode is p dan is:

$$p = \frac{\frac{Ki}{100}(1+\frac{i}{100})^n}{(1+\frac{i}{100})^n-1}$$

of $$K = \frac{100p\left[(1+\frac{i}{100})^n-1\right]}{i(1+\frac{i}{100})^n}$$

of $$n = \frac{\log\left|\frac{p}{p-\frac{Ki}{100}}\right|}{\log(1+\frac{i}{100})}$$

Voorbeeld 6

'n Persoon koop 'n huis en neem 'n bouverenigingverband van R30 000,00, terugbetaalbaar in 30 jaar, teen 18% rente per jaar. Wat sal die maandelikse paaiement wees?

$i = 18\%$ p.j. $= \frac{18}{12}\%$ p.m.

$n = 30 \times 12 = 360.$

$$P = \frac{\frac{(30000)(18)}{100(12)}\left(1 + \frac{18}{(12)(100)}\right)^{360}}{\left(1 + \frac{18}{(12)(100)}\right)^{360} - 1}$$

$= R452,13.$

Voorbeeld 7

As by voorbeeld 6 die persoon bereid is om R500,00 per maand te betaal, hoelank sal dit neem om die verband in voorbeeld 6 af te betaal?

$$n = \frac{\log\left[\frac{500}{500 - \frac{(30000)(18)}{(100)(12)}}\right]}{\log\left(1 + \frac{18}{(12)(100)}\right)}$$

$= 154,6$

≈ 155 maande

$= 12$ jaar en 10 maande.

OPGAWE 17

1. 'n Persoon belê R1 550,00 teen 14% saamgestelde rente, maandeliks saamgestel, vir 5 jaar. Wat sal die eindbedrag van sy belegging wees. Wat was die jaarlikse effektiewe rente op die belegging?

2. Jan koop 'n gewone annuïteit. Hy betaal R50,00 per maand vir 12 jaar, teen 'n rentekoers van 16% per jaar.

   a. Wat sal die eindbedrag wees?
   b. Indien Jan 'n bedrag van R30 000,00 wil hê na 12 jaar, hoe groot moet sy maandleikse betalings wees.

3. 'n Persoon koop 'n motor op huurkoop. Hy leen R5 000,00 teen 18% rente per jaar en is bereid om R200,00 per maand af te betaal. Bereken die aantal betalings wat hy moet maak.

4. R4 500,00 word elke ses maande in 'n vooruitbetaalde aannuïteit teen 10% per jaar belê. Bereken die eindbedrag na vyf jaar.

5. R500,00 word belê vir 5 jaar. As na die 5 jaar die eindbedrag R650,00 was wat was die enkelvoudige rentekoers, per jaar, op die belegging?

6. Hoeveel geld kan 'n persoon leen, teen 15% rente per jaar, as hy bereid is om R250,00 per maand vir 3 jaar te betaal?

7. Hoe lank moet 'n persoon R500,00, kwartaalliks, in 'n gewone annuïteit belê, teen 20% rente per jaar, om 'n eindbedrag van R12 000,00 op te bou?

BYLAE 1

ADDISIONELE OEFENINGE

Die maandelikse bedrag (in rand) wat 'n sekere persoon gedurende 1983 gespaar het is as volg:

62; 29; 46; 87; 21; 50; 38; 29; 10; 73; 65; 43

Bereken die:

(i) modus (Mo)
(ii) rekenkundige gemiddelde ($\bar{x}$)
(iii) mediaan (Me)
(iv) gemiddelde afwyking (G)
(v) 30 ste persentiel (P_{30})

'n Verslaggewer het die volgende gegewens ingesamel in 'n opname wat by verskillende motorhawens gemaak is ten opsigte van die herstelkoste vir 'n bepaalde enjindefek:

R35 R100 R50 R50 R75 R45 R50 R65 R40 R50 R40 R60
R50 R55 R75

Bereken die:

(i) mediaan (Me)
(ii) modus (Mo)
(iii) gemiddelde ($\bar{x}$)
(iv) vierdie desiel (d_4)

'n Motoris ry teen 'n gemiddelde spoed van 80 km/u vanaf punt A na punt B. Daarna draai hy om en ry teen 'n gemiddelde spoed van 120 km/u vanaf punt B na punt A. Bereken die motoris se gemiddelde spoed vir die hele rit.

'n Sekere student het in 'n eksamen 70% en 65% vir Statistiek en Ekonomie respektiewelik behaal. Vergelyk sy prestasie in hierdie twee vakke indien die klasgemiddeldes 56% en 52%; en die standaardafwykings 10% en 8% respektiewelik vir Statistiek en Ekonomie was. In watter vak was sy prestasie die beste? Motiveer u antwoord.

Die gemiddelde jaarlikse salaris wat aan al die werknemers in 'n maatskappy betaal is, was R4 500. Die gemiddelde jaarlikse salaris wat aan manlike en vroulike werknemers van die maatskappy betaal is, was R4 680 en R3 780 onderskeidelik. Bepaal die persentasie manlike en vroulike werknemers wat in diens is by die maatskappy.

Die waarde van maandelikse motorkarverkope (in duisend rand) die afgelope 120 maande moet in 'n frekwensietabel voorgestel word. Die minimum verkope in hierdie periode was 121 en die maksimum 276. Gee 'n geskikte stel klasintervalle vir die data. Toon duidelik alle stappe aan wat u volg.

In 'n studie na die massas van 'n sekere groep beeste is 'n steekproef van 100 beeste geneem. Die maksimum en minimum massa was 210 en 63 kg respektiewelik. Ontwerp 'n geskikte stel klasintervalle vir die data.

8. Die verdeling van die uurlikse lone van werknemers by 'n sekere fabriek is soos volg:

Uurlikse loon	Aantal werknemers
[3,00 ; 4,00)	2
[4,00 ; 5,00)	8
[5,00 ; 6,00)	18
[6,00 ; 7,00)	28
[7,00 ; 8,00)	20
[8,00 ; 9,00)	12
[9,00 ; 10,00)	8
[10,00 ; 11,00)	5

Bereken die:

(i) modus (Mo)
(ii) rekenkundige gemiddelde ($\bar{x}$)
(iii) standaardafwyking (s)
(iv) koëffisiënt van variasie (V)

9. In die volgende tabel word die ouderdomsverdeling van nuwe polishouers onder 'n sekere soort versekering gedurende die jaar 1982 getoon:

Ouderdomsgroep (in jaar)	% Nuwe polishouers in groep
[18 ; 22)	4
[22 ; 26)	8
[26 ; 30)	15
[30 ; 34)	28
[34 ; 38)	20
[38 ; 42)	11
[42 ; 46)	8
[46 ; 50)	5
[50 ; 54)	1

a. (i) Trek 'n histogram en veelhoek van die gegewens.
(ii) Stel die kumulatiewe frekwensieverdelingstabel op en skets die ooreenstemmende veelhoek.
(iii) Bepaal van bg. grafiek die % polishouers met ouderdomme tussen 25 en 75.

b. Bepaal so noukeurig as moontlik vanaf die grafieke wat u alreeds getrek het, die waardes van

(i) die modus
(ii) die mediaan.

Die volgende tabel toon die lengte (in sentimeters) van 'n aantal studente aan 'n sekere Technikon:

Lengte (cm)	Aantal studente
[155 ; 163)	5
[163 ; 171)	14
[171 ; 179)	26
[179 ; 187)	37
[187 ; 195)	32
[195 ; 203)	9
[203 ; 211)	2

a) Trek 'n histogram en frekwensieveelhoek van bostaande inligting op dieselfde asse.

b) Stel 'n kumulatiewe frekwensieverdelingstabel op en trek die ooreenstemmende veelhoek.

c) Gebruik die geskikte grafieke om die volgende te bepaal (dui die waardes duidelik op die grafiek aan):

(i) die aantal studente met 'n lengte van minder as 173 cm.
(ii) die % studente met 'n lengte meer as 184 cm.
(iii) die modus.

Die frekwensieverdeling van die aantal besoekers se binnekomstye by 'n sekere skou vir 'n bepaalde dag is as volg: (Die openingstye van die hekke is om nul uur).

Tyd in uur	Aantal besoekers in 1 000
[0 ; 2)	5
[2 ; 4)	13
[4 ; 6)	27
[6 ; 8)	24
[8 ; 10)	10
[10 ; 12)	5
[12 ; 14)	2

a) Stel die gegewens grafies voor met behulp van 'n histogram.

b) Stel die kumulatiewe frekwensieverdeling op en teken die ooreenstemmende kumulatiewe frekwensieveelhoek.

c) Bepaal, grafies: (Toon elke aflesing aan op grafiek).

(i) Na watter tyd het 50 000 besoekers die hekke binnegegaan?
(ii) Hoeveel besoekers het na 9 uur reeds die skou besoek?
(iii) Op watter tyd het die meeste besoekers die skouterrein binnegegaan (modus)?

d) Bereken vir bogenoemde gegewens die

(i) rekenkundige gemiddelde ($\bar{x}$)
(ii) standaardafwyking (s)
(iii) koëffisiënt van variasie (V), van die tyd wat dit besoekers geneem het om by die skou aan te kom sedert die opening van die hekke.

12. Die frekwensieverdeling van die tye tussen die aankomste van beseerde pasiënte by 'n hospitaal is as volg:

tyd (min.)	0,0 3,0	3,0 6,0	6,0 9,0	9,0 12,0	12,0 15,0	15,0 18,0	18,0 21,0
aantal pasiënte	210	130	75	40	20	15	10

(a) Stel die data vir 'n frekwensie-histogram voor.

(b) Bereken die:

(i) rekenkundige gemiddeld ($\bar{x}$)
(ii) variansie (s^2)
(iii) koëffisiënt van variasie (V)
(iv) mediaan (Me)

13. Die volgende tabel toon die verdeling van pryse van woonhuise wat in 'n sekere stad gedurende Maart 1981 verkoop is:

Prys van woonhuis (in R'000)	Aantal verkoop
[20 ; 25)	55
[25 ; 30)	80
[30 ; 35)	110
[35 ; 40)	240
[40 ; 45)	125
[45 ; 50)	70
[50 ; 55)	50
[55 ; 60)	40
[60 ; 65)	30
TOTAAL	800

(a) Stel 'n kumulatiewe frekwensieverdelingstabel op en trek die ooreenstemmende veelhoek.

(b) Maak gebruik van die veelhoek om die persentasie verkopers te bepaal wat woonhuise met pryse van meer as R52 000 verkoop het.

(c) Bereken die volgende:

(i) die rekenkundige gemiddelde
(ii) die standaardafwyking
(iii) die modus
(iv) die mediaan
(v) 'n maat van skeefheid.

14. Beskou die volgende data:

X	5	8	7	6	10	3	3
Y	10	4	9	10	3	5	8

(a) Bereken die:

(i) korrelasie koëffisiënt (r)
(ii) die vergelyking van die regressielyn.

(b) Stel die lyn op 'n verspreidingsdiagram voor.

(c) Skat die waarde van Y as X = 9.

15. Die Bestuurder van 'n supermark stel belang in die moontlike verband wat daar mag bestaan tussen die weeklikse volume verkope en die bedrag wat spandeer word op advertensieveldtogte. Die volgende gegewens is oor 'n tydperk van ses weke ingesamel:

y: weeklikse volume verkope	10,2	11,5	16,1	20,3	25,6	28
x: Bedrag spandeer op advertensies (R)	100	125	150	200	250	300

a) Bereken die korrelasiekoëffisiënt.
b) Bepaal die regressievergelyking.
c) Skat die weeklikse volume verkope indien 'n bedrag van R220 op advertensies spandeer sou word.

16. Die volgende tabel toon die ouderdomme van vyf mans (x) en die ouderdomme van hul eggenotes (y):

x	48	52	26	29	36
y	43	52	22	25	32

Maak gebruik van die metode van kleinstekwadrate om 'n skatting te maak van 'n vrou se ouderdom indien haar man 42 jaar oud is.

17. Die aantal televisielisensies (in duisend) en die aantal persone wat bioskope besoek het (in miljoen) in 'n sekere land word gegee deur die volgende tabel:

Jaar	Aantal televisielisensies (in duisend)	Aantal besoekers aan bioskope (in miljoen)
1975	22	12
1976	25	12
1977	28	11
1978	31	10
1979	40	8
1980	45	9
1981	58	8
1982	62	8

Bereken die korrelasiekoëffisiënt.

18. Die koste om lint te vervaardig neem af as die hoeveelheid lint wat vervaardig word toeneem. Sekere metings wat op hierdie feit betrekking het, word geïllustreer in die volgende tabel waar x die lengte van die lint is wat vervaardig word, in honderd meter, en y die koste van die lint is, in sent per meter.

x	1	2	3	5	12
y	10	7,5	5,5	5	3,5

i) Bereken (deur die metode van kleinstekwadrate) die waarden van m en k vir die reguitlyn vergelyking $y = mx + k$.

ii) Gebruik die vergelyking om die volgende te bepaal:

1. Die koste per meter van die lint as 1 000 meter vervaardig is, en

2. Die aantal meter lint wat vervaardig moet word om die koste per meter af te bring tot 4,5 sent per meter.

19. Die tabel toon die studentetal (in duisende) aan in sekere universiteit vir die jare 1971 - 1980.

Jaar	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980
Studentetal	0,73	0,97	1,69	1,94	1,90	3,12	2,90	3,51	3,52	4,1

4.1 Teken 'n verspreidingsdiagram.

4.2 Bereken die korrelasie koëffisiënt en verduidelik die betekenis van u antwoord.

4.3 Bereken die vergelyking van die regressielyn wat die data verteenwoordig en teken sy grafiek op u verspreidingsdiagram.

4.4 Skat die studentegetal vir die jaar 1985.

20. In 'n sekere stad is daar 360 biblioteek-assistente waarvan 30 onopgelei is, 180 twee jaar opleiding ontvang het en die res vier jaar opleiding ontvang het. Stel hierdie inligting voor met behulp van:

i) sektorkaart
ii) staafdiagram.

21. Stel 'n staafdiagram van die volgende gegewens op. (Elke firma moet deur 'n afsonderlike staaf verteenwoordig word):

Verkoopsyfers van firmas A, B en C van passasiers- en handelsvoertuie (in duisende eenhede)

	Firma A	Firma B	Firma C
Passasiersmotors	30	25	15
Handelsvoertuie	15	20	5

22. Die onderstaande tabel toon die jaarlikse verbruikersgoedere verkope in 'n sekere land vir 1983.

Soort verbruikersgoedere	Bedrag in miljoen rand
Bederfbaar	320
Nie-bederfbaar	460
Semi-bederfbaar	1 120
Dienste	700
Totaal	2 600

Teken 'n staafdiagram van hierdie gegewens.

23. Op hoeveel maniere kan 'n komitee met 5 lede saamgestel word uit 9 persone?

24. 'n Bak bevat 5 balle. Hoeveel rangskikkings is moontlik indien 3 balle op 'n keer uit die bak getrek word. Volgorde is nie belangrik nie.

25. Bepaal die waarskynlikheid om 'n koning of 'n diamant te kry indien 1 kaart uit 'n pak speelkaarte getrek word.

26. As $P(A) = 1/3$, $P(B) = 3/4$ en $P(A\ B) = 11/12$ bepaal:

a) $P(A \cap B)$
b) $P(A|B)$
c) $P(B|A)$.

27. Koos was die gesin se wasgoed. Die waarskynlikheid dat hy vergeet om beide die bleikmiddel en sagmaakmiddel by te voeg is 1/4. Die waarskynlikheid dat hy vergeet om die bleikmiddel by te voeg is 3/5. As ons weet dat hy vergeet het om die bleikmiddel by te voeg, wat is die waarskynlikheid dat hy ook vergeet het van die sagmaakmiddel.

28. Houer X bevat 5 rooi en 3 blou balle. Houer Y bevat 5 rooi en 4 blou balle. Een bal word uit houer X getrek en sonder om daarna te kyk in houer Y geplaas, waarna 1 bal nou uit houer Y getrek word. Bepaal die waarskynlikheid dat hierdie bal blou is.

29. 'n Houer bevat 50 ratte waarvan 3 defektief is. Indien 'n steekproef van 2 ratte geneem word, wat is die waarskynlikheid dat beide defektief sal wees.

30. Vyf kaarte word sonder teruglegging uit 'n pak van 52 speelkaarte getrek. Bepaal die waarskynlikheid dat:

a) vier van die kaarte konings is
b) vier van die kaarte konings is en een 'n koningin is
c) minstens een van die kaart 'n A is.

31. ŉ Houer bevat 10 wit, 15 swart, 30 rooi en 25 geel balle. Twee balle word agtereenvolgens uit die houer getrek, bepaal die waarskynlikheid dat:

   i) albei balle wit sal wees as terugplasing plaasvind na die eerste trekking.

   ii) 1 bal rooi en 1 bal geel sal wees as geen terugplasing na die eerste trekking gedoen word nie.

32. ŉ Toetsvraestel bestaan uit 20 vrae wat beantwoord moet word deur slegs waar of onwaar te skryf. Wat is die waarskynlikheid vir ŉ student om die korrekte antwoorde van 12 of meer van hierdie vrae te raai?

33. 15% van die artikels wat in ŉ sekere fabriek gemaak word is defektief. Bepaal die waarskynlikheid dat in ŉ steekproef van 30 artikels daar minstens 3 defektief sal wees deur:

   i) die Binomiaalverdeling
   ii) die Poissonbenadering te gebruik.

34. ŉ Marknavorsingsmaatskappy het gevind dat 10% van vraelyste wat uitgestuur word terug kom.
As, vir ŉ sekere studie, 50 vraelyste uitgestuur word wat is die waarskynlikheid dat 7 sal terug kom?

35. Laat X ŉ normaal verdeling besit met gemiddeld = 25 en standaardafwyking = 5. Bereken:

   i) $P(X > 17)$
   ii) $P(19 < X < 21)$

36. Bepaal die waarskynlikheid dat in 200 gooie met ŉ onsydige muntstuk daar tussen 80 en 120 kruise sal wees. Gebruik ŉ normaalbenadering.

37. ŉ Sekere sentrale ontvang gemiddeld 10 uproepe per minuut. Wat is die waarskynlikheid dat daar:

   a) minder as 4
   b) presies 6
   c) tussen 6 en 8 (6 ingesluit, 8 uitgesluit)

   oproepe in ŉ sekere minuut sal inkom?

   (Wenk: Gebruik tabel XII pp 19 van Stoker.)

38. Die massas van ŉ sekere varkras besit ŉ normaalverdeling met $\mu = 25$ en $\sigma^2 = 32$ kg. Wat is die waarskynlikheid om ŉ vark te kry met massa:

   a) kleiner as 21 kg
   b) tussen 23 en 30 kg?

39. Gemiddeld breek daar 1,2 voertuie uit ŉ vloot van 10 voertuie per week. Bereken die waarskynlikheid dat in ŉ sekere week daar presies 2 voertuie sal breek deur

   a) die Poisson-
   b) die Binomiaalverdeling te gebruik.

40. Dit is bekend dat 10% van alle elektriese gloeilampies wat deur ŉ sekere fabriek vervaardig word, defektief is. Wat is die waarskynlikheid dat daar in ŉ ewekansige steekproef van 150 gloeilampies, meer as 25 defektiewes sal wees? (Gebruik die normaalbenadering van die binomiaalverdeling.)

41. Indien ŉ tikster gemiddeld 3 tikfoute per bladsy maak, wat is die waarskynlikheid dat sy ŉ bladsy sal tik met hoogstens 1 tikfout daarop?

42. Die statistiekpunte van 600 studente is normaal verdeel met ŉ gemiddelde van 50 persent en ŉ standaardafwyking van 10 persent. Bereken:

   i) die aantal studente wat tussen 40 en 60 persent behaal het.
   ii) die aantal studente wat tussen 30 en 80 persent behaal het.

43. As die standaardafwyking van die leeftye van televisiebuise op 100 uur geskat word, bepaal wat die grootte van die steekproef moet wees ten einde 95% seker te wees dat die fout in die geskatte gemiddelde leeftyd van televisiebuise nie 20 uur oorskry nie. U mag ŉ normaalverdeling aanvaar.

44. ŉ Masjien maak komponente met ŉ gemiddelde diameter van 20 mm en ŉ standaardafwyking van 0,2 mm.
Watter verhouding van die komponente vervaardig sal ŉ diameter tussen 19,7 en 20,4 mm hê? Aanvaar die data is normaal verdeel.

45. Staalstawe met ŉ vereiste lengte van 100 mm word vervaardig. Die lengtes van die staalstawe is normaal verdeel met ŉ gemiddelde lengte van 100,4 mm en standaardafwyking van 0,8 mm. Dit kos 20c om een staalstaaf te maak en is onmiddellik bruikbaar as die lengte tussen 99 mm en 101 mm is. Stawe wat korter is as 99 mm kan nie gebruik word nie, maar het ŉ afvalwaarde van 5c per staaf. Stawe wat langer as 101 mm is kan korter gemaak word tot ŉ aanvaarbare lengte teen ŉ koste van 8c per staaf.

   Bepaal die gemiddelde koste om ŉ staaf van bruikbare lengte te maak.

46. ŉ Masjien word aangedryf deur 3 batterye van dieselfde tipe. Die masjien sal aanhou werk as ten minste twee batterye in werkende toestand is. Die waarskynlikheid dat ŉ battery gedurende die eerste 10 uur sal ingee is 0,2.

   Bepaal die waarskynlikheid dat die masjien aanhoudend vir 10 uur sal werk.

47. 150 elektriese gloeilampe in 'n steekproef het 'n gemiddelde leeftyd van 1 400 uur en 'n standaardafwyking van 120 uur gehad. Bepaal die 95% betroubaarheidsinterval vir die ware gemiddelde en -variansie van die leeftyd van al die gloeilampe.

48. 'n Ewekansige steekproef van 150 persone in Pretoria het getoon dat 127 gereeld na sportprogramme op televisie kyk. Toets of die populasieverhouding sportkykers 0,80 is ($\alpha = 0,05$).

49 'n Fabriek gebruik twee verskillende prosesse om elektriese gloeilampies te vervaardig. Ten einde vas te stel of daar 'n betekenisvolle verskil in die gehalte van die gloeilampies is wat deur die twee prosesse vervaardig word, word ewekansige steekproewe uit elke vervaardigingsproses getrek en die brandleeftye van die gloeilampies word bepaal. Die volgende resultate is verkry:

	Steekproefgrootte	Gemiddelde brandleeftyd	Standaardafwyking
Proses I	$n_1 = 11$	$\bar{x}_1 = 808,8$ uur	$s_1 = 9,5$ uur
Proses II	$n_2 = 14$	$\bar{x}_2 = 824,4$ uur	$s_2 = 10,7$ uur

Aanvaar dat die brandleeftye van elektriese gloeilampies normaal verdeel is en toets of proses II 'n beter gehalte gloeilampie lewer as proses I ($\alpha = 0,05$).

50. Die volgende twee steekproewe uit normaal verdeelde populasies word gegee:

Steekproef 1: 11,9; 17,3; 15,6; 24,9; 16,9; 20,3; 20,6

Steekproef 2: 17,9; 18,6; 20,8; 11,2; 16,5; 24,8; 24,2; 28,0

Bepaal of die twee populasiegemiddeldes betekenisvol verskil. Betekenispeil = 5%.

51. Vir twee steekproewe word die volgende gegee:

	Grootte	Rekenkundige Gemiddelde	Standaardafwyking
Steekproef 1	$n_1 = 8$	$\bar{x}_1 = 28,7$	$s_1 = 2,8$
Steekproef 2	$n_2 = 10$	$\bar{x}_2 = 27,3$	$s_2 = 3,2$

Bepaal of die twee populasiegemiddeldes betekenisvol verskil by die 1% betekenispeil.

52. In 'n steekproef van 15 elemente was gevind dat $\bar{x} = 15,7$ en $s = 2,36$. Gee 'n 99% betroubaarheidinterval vir die populasiegemiddeld, μ en die populasievariansie, σ^2.

53. Om die gelykheid van twee verhoudings te toets is die volgende waargeneem:

$n_1 = 350$; $X_1 = 140$; $n_2 = 480$; $X_2 = 218$.

Toets by $\alpha = 0,01$ of die twee verhoudings uit identiese populasie kom.

54. Die gemiddelde deursnit van wolvesels word gegee as 26,5 mikron. In 'n ontleding van 20 vesels was die gemiddelde deursnit 28 mikron met 'n standaardafwyking van 2,5 mikron.

Is hierdie verskille betekenisvol by $\alpha = 0,01$?

55. Die standaardafwyking van die breeksterkte van sekere kabels vervaardig deur 'n sekere firma word aangegee as 240 kN. Nadat 'n verandering in die vervaardigingsproses aangebring is het 'n steekproef van 8 kabels 'n standaardafwyking van 300 kN getoon. Is hierdie skynbare toename in die standaardafwyking betekenisvol by die 5% betekenispeil?

56. Ewekansige steekproewe van 300 skroewe vervaardig deur masjien A en 200 skroewe vervaardig deur masjien B het 24 en 12 defektiewe skroewe respektiewelik getoon. Bepaal of daar 'n betekenisvolle verskil tussen die werking van die twee masjiene is by die 1% betekenispeil.

57. Die variansie van die wanddikte van 'n sekere soort staalpyp moet 2,5 mm wees. In 'n ondersoek van 25 pype is 'n variansie van 1,58 mm waargeneem. Toets by $\alpha = 0,05$ of die pype nog aanvaarbaar is.

(Wenk: Doen 'n tweekantige toets.)

58. Die gemiddelde inhoud van blikkies ingelegde groente moet 610 g wees. In 'n ondersoek van 8 blikkies is die volgende massas waargeneem:

620,5; 600,9; 599,1; 624,7; 598,7; 608,0; 611,8; 596,6 g.

Kan ons aanvaar dat die inhoud van die blikkies nog voldoende is by $\alpha = 0,05$?

59. Die volgende twee steekproewe kom uit normaal verdeelde populasies.

steekproef						
A	7	8	13	19	24	
B	11	14	6	23	17	9

Toets by $\alpha = 0,05$ of die twee steekproefgemiddeldes gelyk is.

60. Bepaal of die aannames van normaalverdeling in oefeninge 50, 58 en 59 geregverdig was.

61. Vir twee steekproewe uit normaalverdeelde populasies word die volgende gegee:

	Grootte	Rekenkundige Gemiddelde	Variansie
Steekproef 1	$n_1 = 100$	$\bar{x}_1 = 25,7$	$s_1^2 = 7,84$
Steekproef 2	$n_2 = 120$	$\bar{x}_2 = 26,9$	$s_2^2 = 11,56$

Bepaal of die twee populasiegemiddeldes betekenisvol verskil. Betekenispeil = 5%.

62. In 360 gooie met twee seskantige dobbelstene is die volgende waargeneem:

58 agts
35 tiene
18 elfs.

Bepaal of die dobbelstene onsydig is. Betekenispeil = 5%.

63. Ondersoek die verspreiding van die residue en toets of die korrelasiekoëffisiënte van oefening 14 tot 19 betekenisvol is. $\alpha = 0,05$.

64. Die volgende 3 steekproewe word gegee:

steekproef 1	2	3
90	84	48
88	83	64
92	85	74
86	84	69

Voer 'n eenrigting analise van variansie uit om te bepaal of die steekproefgemiddeldes verskil en indien wel bepaal presies waar lê die verskille. $\alpha = 0,05$.

65. Die volgende resultate is gedurende 'n eksperiment bepaal:

LESINGS

		A	B	C	D
Persone	1	16,3	22,6	21,7	18,7
	2	18,8	23,9	22,8	19,3
	3	19,4	24,1	23,9	17,6
	4	17,5	22,1	22,6	15,8
	5	19,6	22,7	23,0	20,9

Deur 'n variansie-analise uit te voer, bepaal of die resultate van die persone en die stelle lesings betekenisvol verskil en indien wel, bepaal watter stelle verskil betekenisvol. Betekenispeil = 5%.

66. Die volgende vier stelle lesings is gedurende 'n eksperiment bepaal:

LESINGS

A	B	C	D
16,3	20,5	20,3	21,7
18,4	19,6	20,6	20,3
17,6	18,3	19,6	19,8
17,3	19,5	18,5	20,9
16,7	20,1	19,9	21,2

Deur 'n variansie-analise uit te voer bepaal of die vier stelle lesings betekenisvol verskil en indien wel bepaal watter stelle verskil betekenisvol.
Betekenispeil = 5%.

(Gegee: Die totale som van die vierkante = 43,57. Die som van die vierkante vir tussen die stelle = 33,30.)

67. Drie verskillende vervaardigingsprosesse word gebruik om 'n sekere tipe plastiek te vervaardig. Ewekansige steekproewe word uit elke vervaardigingsproses getrek en die elastisiteit van die monsters word bepaal. Die volgende resultate is verkry:

Proses A	Proses B	Proses C
4,2	5,6	3,2
1,1	5,1	2,5
3,7	4,4	2,9
2,6	4,2	3,6
2,1	4,2	3,2
3,7	5,1	4,1

Toets of die gemiddelde elastisiteit van plastiek vervaardig deur die drie prosesse betekenisvol van mekaar verskil, en indien wel, bepaal dan watter prosesse dit is wat betekenisvol van mekaar verskil. $(\alpha = 0,05)$.

68. Die volgende twee steekproewe word gegee.

Steekproef A: 8; 15; 23; 17; 35; 14; 9

Steekproef B: 22; 11; 16; 18; 27; 30; 39; 24

Dit is nie bekend of die steekproewe normaal verdeel is nie. Bepaal of hulle afkomstig is uit identiese populasies. Betekenispeil = 5%.

69. Om te bepaal of 'n sekere behandeling beter is as 'n kontrole is 'n eksperiment uitgevoer en die volgende waargeneem:

Behandeling	26	22	14	29	23	18	25
Kontrole	15	21	24	10	16	19	

Deur 'n Mann-Whitney toets uit te voer bepaal of die behandeling wel beter gevaar het by $\alpha = 0{,}01$.

70. Die volgende twee steekproewe word gegee. Dit is nie bekend of hulle normaalverdeel is nie.

Steekproef A

8,6; 7,8; 8,3; 6,6; 10,5; 6,3; 9,3; 8,4; 9,7.

Steekproef B

6,5; 10,0; 7,0; 9,8; 8,5; 9,2; 9,0; 8,2; 10,8; 8,7

Deur die Mann-Whitney-toets te gebruik, bepaal of daar 'n betekenisvolle verskil tussen die twee steekproewe is. Betekenispeil = 5%.

71. 'n Houer bevat 'n groot aantal balle van vier verskillende kleure naamlik rooi, oranje, blou en geel. 'n Gelykkansige steekproef van 48 balle het 15 rooi, 11 oranje, 14 blou en 8 geel balle getoon. Toets die hipotese dat die houer gelyke verhoudings van die verskillende kleure balle bevat. Betekenispeil = 10%.

72. In sekere genetiese kruisings word die tipes, voorgestel deur PQ, Pq, pQ en pq, verwag om voor te kom in die verhouding 9 : 3 : 3 : 1. Vir 'n sekere eksperiment was die waargenome frekwensie (f) as volg:

Tipe	PQ	Pq	pQ	pq	Totaal
f	253	93	87	47	480

Bevredig hierdie waardes die gegewe verhouding? Betekenispeil = 1%.

73. Die volgende 3 x 3 gebeurlikheidstabel word gegee:

	Steekproef 1	Steekproef 2	Totale
Klasse 1	18	25	43
2	24	33	57
3	37	38	75
Totale	79	96	175

Bepaal of die twee steekproewe betekenisvol verskil: Betekenispeil = 5%.

74. Die volgende 2 x 2 gebeurlikheidstabel word gegee:

	kolom 1	kolom 2
ry 1	8	10
ry 2	13	7

Is die rye en kolomme onafhanklik by $\alpha = 0{,}01$?

75. 'n Meningspeiling onder studente aan 'n sekere Technikon ten op- sigte van 'n voorstel dat klasgelde verhoog word, het die volg resultate opgelewer:

Geslag van student	Mening: Daarvoor	Daarteen	Weet nie
Manlik	59	69	14
Vroulik	91	54	13

Gebruik 'n χ^2 -toets om in hierdie 2 x 3 gebeurlikheidstabel, te toets of die menings van manlike en vroulike studente betekenisvol van mekaar verskil. ($\alpha = 0{,}05$).

76. Beskou die data:

Artikel	Gewig	Prys per artikel 1975	Prys per artikel 1982	Hoeveelheid verkoop (in duis 1975	Hoeveelheid verkoop 1982
A	12	10c	12c	15	18
B	6	50c	60c	10	8
C	5	100c	90c	3	3
D	7	30c	36c	20	30

Gebruik 1975 as basisjaar en bereken die volgende indekse vi

i) Laspeyres se prysindeks
ii) Paasche se prysindeks
iii) Fisher se ideale prysindeks.

77. Die kleinhandelspryse van 3 basiese verbruikersartikels en die hoeveelhede verkoop vir 1975 en 1983 is soos volg:

Artikel	Pryse (sent per eenheid)		Hoeveelhede		Gewig
	1975	1983	1975	1983	
Brood	22	36	60	75	10
Vleis	175	283	80	100	12
Botter	95	215	45	50	6

Bereken die volgende indekse vir 1983 met 1975 = 100:

i) enkelvoudige prysindeks vir brood
ii) geweegde saamgestelde prysindeks
iii) Paasche se prysindeks
iv) Laspeyres se prysindeks
v) Fischer se prysindeks.

78.

	1981		1982	
	Prys (in R)	Hoeveelheid	Prys (in R)	Hoeveelheid
Kaas	3,20	36	3,68	41
Melk	0,40	730	0,51	700
Eiers	0,90	90	1,00	96

Bereken en interpreteer (met 1981 as basisjaar) die volgende:

i) Laspeyres se prysindeks.
ii) Paasche se prysindeks.
iii) Die geweegde gemiddelde prysindeks. Gebruik gewigte van 1; 20 en 3 vir kaas, melk en eiers respektiewelik.

79.

Item	Prys in sent		Produksie (kg)	
	1979	1982	1979	1982
Produk A	100	170	1 200	3 200
Produk B	20	30	15 000	20 000
Produk C	40	50	8 000	10 000

Bereken met 1979 as basisjaar, vir 1982:

i) die Laspeyres prysindeksyfer
ii) die Paasche prysindeksyfer
iii) Fischer se prysindeksyfer
iv) Drobisch se prysindeksyfer.

80. Die volgende getalle stel die kwartaallikse winste (in miljoen R) van ŉ sekere nywerheid gedurende die afgelope 4 jaar voor.

Jaar	1979				1980				1981				1982			
Kwartaal	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Wins (R miljoene)	6	10	18	7	4	12	20	9	2	6	18	11	8	8	24	13

Bereken:

i) die gesentreerde vier-kwartaallikse bewegende gemiddelde.
ii) die seisoensindeks.

81. Die volgende tabel toon die kwartaallikse verkope (in R'000) van ŉ kleinhandelaar aan:

		1979	1980	1981	1982
Kwartaal	1	72	74	77	81
	2	57	56	60	65
	3	48	50	52	57
	4	77	78	79	82

i) Bereken die tipiese seisoenindekse.
ii) Indien die verwagte verkope vir 1983 R305 000 is, bereken die verwagte verkope vir die derde kwartaal van 1983.

82. ŉ Vader belê by die geboorte van sy seun R1 000 in ŉ spaarrekening teen 0,08 p.e.p.j. saamgestelde rente, ten einde vir sy seun se moontlike studie aan ŉ universiteit voorsiening te maak. Oor hoeveel sal hy beskik as die seun op 19-jarige ouderdom na ŉ universiteit gaan?

83. i) Hoe lank sal dit R1 250 neem om aan te groei tot R4 167 teen 0,07 p.e.p.j., halfjaarliks bygereken?

ii) As R1 250 in 14 jaar aangroei tot R2 475, teen watter rentekoers p.j. geskied dit?

84. Bepaal die slotbedrag na 6 jaar van R2 364,50 teen 9% p.j., kwartaalliks saamgestel.

ŉ Persoon het R500 wat hy vir 6 jaar kan belê. Hy kan dit óf teen 6% p.j. enkelvoudige rente óf teen 5% p.j. samegestelde rente, kwartaalliks bygereken, belê. Watter een van die twee beleggings is met hoeveel die voordeligste?

85. ŉ Persoon belê R200 teen 8% p.j., halfjaarliks saamgestel. Vind hoeveel rente (i) in die vyfde jaar; (ii) in die tiende jaar bygereken word.

BYLAE 2

DATASTEL

HIERDIE DATASTEL IS UIT 'N GELYKMATIGE-VERDELING GEGENIREER.

DIE DOEL HIERVAN IS OM DIT AS 'N POPULASIE TE BESKOU, VERSKILLENDE STEEKPROEWE HIERUIT TE TREK EN DIE DATA TE ONTLEED.

NO	X	Y	NO	X	Y
1	38.7	1.16	51	89.9	9.5
2	75.4	9.54	52	95.9	1.95
3	67.9	3.78	53	59.5	3.13
4	73.3	6.25	54	94.8	9.71
5	55.8	7.34	55	34.5	4.34
6	52.4	6.66	56	90.7	1.1
7	35.2	1.28	57	54.7	3.28
8	59.1	8.09	58	91.1	7.57
9	35.9	3.8	59	45.8	5.74
10	88.1	4.25	60	51.5	8.79
11	83.8	7.59	61	90.6	4.01
12	61.4	8.45	62	95.4	8.13
13	78.2	3.73	63	92.8	2.78
14	41.5	6.93	64	86.8	1.17
15	73.2	3.12	65	89.7	5.28
16	89.6	3.04	66	90.3	8.1
17	41.3	1.55	67	55.6	1.54
18	85.2	1.01	68	43.2	9.78
19	37.5	7.12	69	38.6	2.89
20	57.5	5.46	70	34.5	6.88
21	80.5	1.5	71	71.4	8.53
22	61.5	3.64	72	46.8	3.51
23	55.1	3.25	73	97.1	6.03
24	52.5	8.34	74	47.8	1.85
25	77.5	4.18	75	33.3	6.19
26	95.7	8.07	76	56.6	4.46
27	56.1	1.95	77	66.5	1.76
28	91.3	1.35	78	36.5	4.78
29	98.8	9.95	79	95.7	4.67
30	30.8	7.94	80	88.9	1.68
31	33.6	4.94	81	71.3	9.73
32	73.9	8.16	82	84.7	3.63
33	80.2	5.26	83	48.6	7.52
34	45.5	8.82	84	60.5	5.86
35	96.9	4.65	85	36.3	8.53
36	34.9	2.86	86	38.1	8.85
37	62.2	5.72	87	72.4	7.09
38	85.1	7.17	88	94.5	4.69
39	56.1	2.39	89	39.1	6.65
40	78.9	8.5	90	94.5	5.74
41	37.5	2.29	91	89.1	8.43
42	87.9	5	92	82.7	1.48
43	50.3	4.64	93	38.8	1.65
44	47.4	9.1	94	44.8	4.88
45	70.6	3.12	95	41.1	8.66
46	58.1	3.99	96	60.1	3.96
47	56.1	5.71	97	33.6	6.87
48	38.8	7.12	98	96.1	6.06
49	32.2	1.64	99	42.1	7.33
50	94.8	5.37	100	41.7	8.89

NO	X	Y	NO	X	Y
101	60.3	4.46	151	77.2	1.94
102	59.9	6.19	152	61.1	9.9
103	96.6	1.21	153	70.7	3.96
104	75.9	2.51	154	40.8	3.4
105	80.4	7.34	155	31.9	2.85
106	86.9	3.02	156	31.4	7.69
107	61.6	4.91	157	34.1	7.19
108	95.1	6.52	158	60.6	1.91
109	68.7	1.48	159	41.1	6.97
110	65.8	7.4	160	84.6	7.98
111	80.6	4.43	161	56.3	1.47
112	68.3	3.13	162	80.5	2.84
113	96.6	3.32	163	34.1	1.11
114	72.5	4.62	164	82.2	1.46
115	77.9	5.91	165	63.7	8.33
116	74.3	9.93	166	61.7	6.83
117	30.8	6.2	167	41.1	6.44
118	47.6	3.38	168	95.6	7.81
119	70.5	9.66	169	57.6	7.5
120	87.4	7.27	170	84.9	3.47
121	93.8	9.74	171	62.9	2.8
122	84.9	6.24	172	40.4	1.71
123	54.7	8.19	173	93.5	1.68
124	85.5	7.36	174	93.4	5.25
125	42.2	3.39	175	75.8	7.32
126	36.3	9.24	176	74.1	2.05
127	83.3	1.93	177	77.3	9.29
128	92.6	1.39	178	98.4	9.68
129	33.7	9.29	179	30.7	3.22
130	48.8	8.8	180	48.8	7.91
131	59.1	1.83	181	71.2	5.47
132	54.6	1.02	182	37.6	3.91
133	83.7	6.4	183	66.6	6.88
134	60.1	1.65	184	31.7	1.93
135	53.3	7.3	185	79.8	9
136	73.4	1.46	186	62.5	2.9
137	83.7	2.34	187	47.1	1.57
138	35.7	8.21	188	90.4	3.31
139	41.3	7.95	189	91.7	7.19
140	63.3	8.52	190	35.8	7.21
141	81.1	8.54	191	56.1	1.75
142	81.8	2.45	192	49.7	5.86
143	80.6	6.4	193	91.4	2.42
144	85.7	3.03	194	82.3	7.15
145	81.4	6.08	195	87.3	2.51
146	41.9	1.69	196	89.1	4.19
147	46.3	1.24	197	71.4	1.2
148	84.9	8.16	198	50.4	4.73
149	66.8	8.95	199	99.2	9.11
150	71.1	3.62	200	43.4	1.9

NO	X	Y	NO	X	Y
201	83.2	2.24	251	74.5	8.16
202	35.8	5.51	252	47.6	6.94
203	79.4	8.61	253	37.8	7.58
204	35.6	7.09	254	49.7	4.63
205	40.9	1.62	255	44.3	6.07
206	43.5	8.82	256	56.1	8.16
207	70.2	3.55	257	33.4	8.74
208	91.9	7.85	258	73.2	1.22
209	47.5	2.75	259	58.1	5.48
210	36.4	9.99	260	80.9	1.61
211	90.1	6.25	261	79.4	5.07
212	48.5	1.47	262	63.1	9.22
213	79.2	1.84	263	91.4	1.52
214	66.8	6.14	264	82.1	2.36
215	89.6	6.31	265	70.1	9.79
216	74.1	8.92	266	93.9	6.59
217	43.2	2.8	267	99.5	2.84
218	43.1	5.57	268	90.1	7.57
219	94.9	9.8	269	67.2	3.23
220	80.3	1.1	270	75.2	4.35
221	94.2	6.79	271	71.1	7.51
222	85.9	5.43	272	92.1	7.03
223	76.8	4.36	273	71.9	6.37
224	75.2	9.77	274	56.1	7.98
225	82.9	1.86	275	80.2	8.96
226	80.9	1.35	276	59.6	3.31
227	67.1	9.27	277	62.1	5.16
228	75.2	2.32	278	65.4	2.64
229	45.4	4.39	279	59.6	4.45
230	71.1	5.24	280	77.7	4.17
231	91.1	6.74	281	35.9	5.68
232	79.1	5.18	282	56.3	5.53
233	50.1	7.79	283	76.8	2.39
234	82.1	3	284	35.4	9.65
235	44.1	3.31	285	79.2	2.42
236	99.8	7.24	286	93.3	7.2
237	79.8	7.95	287	42.5	1.1
238	44.2	5.83	288	62.6	3.02
239	39.8	7.55	289	49.4	1.09
240	30.1	6.43	290	80.7	5.43
241	36.3	3.67	291	77.6	5.5
242	73.1	6.37	292	74.4	3.59
243	31.2	9.17	293	90.8	1.5
244	55.2	5.78	294	99.2	1.59
245	56.5	9.59	295	30.7	6.09
246	81.4	2.63	296	75.6	8.99
247	98.3	4.1	297	30.9	3.22
248	62.5	2.84	298	44.1	8.02
249	93.2	4.12	299	35.9	4.74
250	77.6	2.53	300	69.4	8.8

NO	X	Y	NO	X	Y
301	71.3	1.76	351	43.6	6.64
302	33.1	7.26	352	94.6	5.8
303	41.4	7.08	353	83.4	7.56
304	53.8	2.55	354	31.1	4.62
305	83.3	1.96	355	97.6	2.81
306	79.5	8.65	356	97.3	9.8
307	76.7	7.18	357	58.6	2.86
308	46.4	8.28	358	79.8	1.74
309	70.7	1.33	359	68.4	6.74
310	70.8	1.01	360	50.5	1.82
311	90.6	3.32	361	44.6	7.33
312	63.1	4.59	362	70.1	4.66
313	96.4	7.06	363	91.3	4.1
314	91.1	1.91	364	76.2	2.48
315	63.1	1.12	365	84.3	7.24
316	30.9	9.25	366	95.8	1.51
317	62.7	1.32	367	69.7	5.17
318	85.6	1.42	368	56.3	2.96
319	77.2	3.89	369	39.8	4.81
320	89.2	7.17	370	94.1	1.94
321	56.4	4.39	371	44.7	5.12
322	47.4	1.51	372	63.3	9.78
323	52.7	7.01	373	97.7	1.98
324	47.6	9.79	374	60.7	1.45
325	47.5	8.1	375	45.8	4.25
326	90.6	4.42	376	54.1	2.78
327	71.5	9.38	377	67.1	3.29
328	96.8	1.38	378	64.7	7.49
329	93.2	4.69	379	81.6	7.46
330	69.2	1.16	380	57.3	1.74
331	52.9	6.54	381	75.6	4.37
332	94.6	1.39	382	70.9	1.65
333	85.1	8.98	383	74.3	8.63
334	77.4	8.81	384	33.7	4.91
335	65.2	1.6	385	43.9	3.56
336	37.8	4.93	386	93.8	4.76
337	38.3	6.03	387	82.2	1.56
338	87.6	9.96	388	77.6	1.26
339	71.3	1.34	389	33.3	8.8
340	78.6	9.54	390	51.3	4.65
341	83.1	7.16	391	37.1	5.42
342	50.5	4.38	392	83.3	7.26
343	74.1	8.75	393	57.8	9.28
344	33.6	4.61	394	55.1	3
345	45.5	3.17	395	50.7	8.94
346	78.1	2.89	396	51.7	3.05
347	99.4	5.25	397	35.6	7.23
348	46.4	4.93	398	59.6	4.91
349	58.3	4.02	399	41.5	7.64
350	93.1	2.97	400	67.5	6.12

NO	X	Y	NO	X	Y
401	71.1	3.62	451	43.4	1.9
402	37.4	1.36	452	61.1	1.3
403	53.5	8.88	453	92.2	8.28
404	89.4	3.6	454	34.1	1.46
405	75.1	1.93	455	85.5	4.54
406	75.6	1.76	456	47.7	1.37
407	37.3	2.76	457	90.8	7.78
408	75.3	2.12	458	42.8	5.73
409	86.9	2.28	459	56.9	1.64
410	81.9	4.77	460	41.1	4.1
411	73.7	2.55	461	59.6	3.85
412	43.4	1.01	462	87.6	3.64
413	31.9	4.06	463	40.2	3.21
414	45.3	2.49	464	85.8	8.56
415	59.1	9.04	465	50.1	2.69
416	40.9	4.76	466	77.1	3.73
417	31.7	8.25	467	53.7	2.05
418	52.1	1.78	468	99.5	8.66
419	41.7	9.17	469	73.7	9.19
420	44.7	1.54	470	49.7	1.69
421	30.6	2.82	471	77.8	1.22
422	50.8	5.15	472	85.2	1.58
423	60.4	6.9	473	90.7	8.75
424	45.5	2.46	474	87.1	6.11
425	48.5	8.53	475	32.8	7.53
426	88.1	4.55	476	72.4	2.46
427	30.1	4.09	477	97.9	5.95
428	66.6	9.47	478	91.2	8.93
429	76.8	3.65	479	68.5	3.06
430	92.9	8.93	480	41.2	5.82
431	55.3	6.92	481	73.3	1.59
432	47.4	2.8	482	58.6	3.5
433	81.7	4.16	483	46.9	2.72
434	65.1	7.21	484	76.8	7.99
435	92.8	5.95	485	92.7	1.37
436	99.8	8.69	486	50.1	7.37
437	48.5	2.97	487	53.5	3.73
438	38.1	9.96	488	87.6	2.6
439	98.1	6.42	489	54.6	2.29
440	73.2	1.24	490	33.4	4.44
441	80.7	7.6	491	52.8	1.85
442	34.4	1.45	492	50.5	7.28
443	98.7	6.22	493	31.6	5.57
444	98.1	9.54	494	94.4	2.81
445	44.4	8.24	495	54.3	6.85
446	57.1	7.32	496	67.6	8
447	34.5	1.48	497	85.6	2.23
448	44.3	3.93	498	65.2	1.42
449	81.4	3.93	499	83.1	1.02
450	34.9	6.65	500	57.1	9.85

BYLAE 3

ANTWOORDE VAN OPGAWES.

1: Bladsy 3

4. a. 109
   b. 676
   c. -9
   d. -35

5. a. 190,32
   b. 382,88
   c. 2756,25
   d. 1186,5

3: Bladsy 15

1. i. 151,25
   ii. 146
   iii. 148
   iv. 47
   v. 11,29
   vi. 190,2 13,79
   vii. 9,1

2. V_A = 33,85 V_B = 23,3

3. i. 11,73   11,9   11,9
   ii. 6,1   3,57   16,11

4. i. 47,44   167,52
   56,31   487,32
   36,91   37,04

   ii. 47,70   294,53

6. 150,67   150,08
   11,59   11,46

5: Bladsy 32

1.	146,5	145,0	145,9	139,8	0,11	-0,0012
	27,2	22,0	25,0	19,9	0,45	0,0016
	4,6	2,52	3,9	2,44	0,67	0,017
	56,1	48,0	55	47,44	0,633	0,031
	51,9	55,3	53,3	48,2	-0,382	0,003

2. a. 48,1
   b. 44,04
   c. 19,15
   d. 0,21   -0,007

3. i. 79,67
   ii. 6,657
   iii. 8,4
   iv. 79,38
   v. 74,58   86,0

6: Bladsy 37

1. a. $\hat{y}$ = -0,07 + 3,09x   r = 0,997
   b. $\hat{y}$ = 3,45 + 3,06x   r = 0,999
   c. $\hat{y}$ = 20 + 2x   r = 0,909

2. 34   50

3. 0,986

4. a. $\hat{y}$ = 0,086 + 0,033x
   c. R3,39

7: Bladsy 45

1. i. 17160.
ii. 3024

3. 1024.

5. 2598960.

7. 0,12.

9. 0,83.

11. i. 0,8.
ii. 0,7.
iii. 0,5.

13. 0,104.

15. 0,87229

17. i. $\frac{14}{500}$

ii. $\frac{290}{500}$

iii. $\frac{163}{210}$

iv. $\frac{10}{14}$

19. i. 0,1540
ii. 0,2338

2. i. 40320.
ii. 10080.

4. 9261000.

6. 600

8. i. $\frac{1}{22}$.

ii. $\frac{3}{11}$.

10. i. $\frac{777}{841}$.

ii. $\frac{378}{406}$

12. i. 0,00198.
ii. 0,00039.

14. $\frac{1}{7}$.

16. I a. 0,0006.
b. 0,0588
II a. 0,0002
b. 0,0196

18. i. 0,0320

ii. 0,75

8: Bladsy 60

1. a. 0,0576 b. 0,5518 c. 0,1941

2. a. 0,0167 b. 0,5859 d. 0,2072

3. a. 0,0708 b. 0,1151

4. a. 0,1369 b. 0,4114 c. 0,5886

5. a. 0,1954 b. 0,9786 c. 0,1106

6. a. 0,15534 b. 0,0162 c. 0,00016

7. a. 0,0329 b. 0,8037

8. a. 0,0378 b. 0,5667 c. 0,9707.

9. 0,00621 b. 0,04786

9: Bladsy 67

1. 90% : (30,02 ; 32,38) (2,907 ; 12,5003)

95% : (29,75 ; 32,65) (2,609 ; 14,99)

2. 95% : (0,09 ; 0,209)
99% : (0,072 ; 0,228)

3. a. $\mu = 2,88$ $\sigma^2 = 0,287$

b. (2,468 ; 3,292) (0,13 ; 1,05)

4. (60,66 ; 90,20) (74,70 ; 94,16)

5. (0,545 ; 0,655)

6. 90% : (960,56 ; 985,84)

99% : (953,4 ; 993,00)

10: Bladsy 81

2. $Z_s = -5,506 < -1,645$ Ho aanvaar

3. $t_s = 3,853 > 1,734$ Ho verwerp

4. $0,26 < F_s = 2,92 < 4,3$ Ho aanvaar

6. $Z_s = -8,97 < -2,576$ Ho verwerp

7. $t_s = -6 < -1,86$ Ho verwerp

8. $Z_s = 1,048 < 1,282$ Ho aanvaar

9. $Z_s = 1,09 < 1,645$ Ho aanvaar

10. $t_s = 1,49 < 2,539$ Ho aanvaar

11. $X^2_s = 21,85 < 30,144$ Ho aanvaar

12. $Z_s = 7,35 > 2,326$ Ho verwerp

13. $t_s = 1,845 > 1,383$ Ho verwerp

14. b. $0,167 < F_s = 2,746 < 6,98$ Ho aanvaar
c. $t_s = 0,984 < 2,201$ Ho aanvaar.

11: Bladsy 88

1. $R_s = 0,997 > 0,7293$ Ho verwerp.
   $R_s = 0,999 > 0,7293$ Ho verwerp.
   $R_s = 0,909 > 0,7293$ Ho verwerp.

2. a. $\hat{y} = 0,925 + 0,020x$.
   $r = 0,099$.
   c. $R_s = 0,099 < 0,6581$ Ho aanvaar.

12: Bladsy 100

1. $F_w = \frac{455,4/5}{10090,8/24} = 0,217 < 2,62$ Ho aanvaar.

2. $F_w = \frac{455,4/5}{3080,8/20} = 0,591 < 2,71$ Ho aanvaar.
   $F_d = \frac{7010/4}{3080,8/20} = 11,377 > 2,87$ Ho verwerp.

3. $F_a = \frac{1078,8/3}{11629,0/12} = 0,371 < 5,95$ Ho aanvaar.
   $F_b = \frac{44677,3/2}{11629,0/12} = 23,051 > 6,93$ Ho verwerp.
   $F_c = \frac{2982,7/6}{11629,0/12} = 0,513 < 4,82$ Ho aanvaar.

4. $F_s = \frac{8647,5/4}{1289,7/16} = 26,820 > 3,73$ Ho verwerp.

13: Bladsy 107

1. $X_s^2 = 22,24 > 7,815$ Ho verwerp.

2. $X_s^2 = 2 < 6,251$ Ho aanvaar.

3. $X_s^2 = 6,12 < 12,592$ Ho aanvaar.

4. $X_s^2 = 0,3978 < 9,210$ Ho aanvaar.

14: Bladsy 112

1. a. $T_s = 6 > 2$ Ho verwerp.
   b. $T_s = 4 > 0$ Ho verwerp.
   c. $T_s = 1 > 0$ Ho verwerp.

2. $W_s = 25,5 > 16$ Ho verwerp.

3. $W_s = 15 > 8$ Ho verwerp.

15: Bladsy 117

1. a. 1982 : 113 122 142
   1983 : 72 165 163
   b. 99 94
   c. 127

2. 107 106 106 107

3. 1982 Basisjaar: 100 109 108 114 119 116 124 129 138 142 133 119 126

16: Bladsy 127

b. $\hat{y} = 53,827 + 0,356t$.

Vooruitskattings

Jaar		t	L_i	S	$LS = \hat{y}$
	I	31	64,86	1,44	93,40
1960	II	33	65,58	0,83	54,43
	III	35	66,29	0,73	48,39
1961	I	37	67,00	1,44	96,48

17: Bladsy 135

1. R3108,70 20,11%
2. a. R21 506,12
   b. R69,75 p.m.
3. 32 maande
4. R59430,54
5. 6% p.j.
6. R7 211,82
7. 4,25 jaar